AFRL-RV-HA-TR-2008-1121

Improving Magnitude Detection Thresholds Using Multi-Station, Multi-Event, and Multi-Phase Methods

D. Schaff
F. Waldhauser

Columbia University
254 Engineering Terrace
1210 Amsterdam Avenue; MC 2205
New York, NY 10027-7003

Final Report

31 July 2008

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
29 Randolph Road
AIR FORCE MATERIEL COMMAND
Hanscom AFB, MA 01731-3010

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release and is available to the general public, including foreign nationals. Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC) (http://www.dtic.mil). All others should apply to the National Technical Information Service.

AFRL-RV-HA-TR-2008-1121 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//Signature//	//Signature//
ROBERT RAISTRICK Contract Manager	PAUL TRACY, Acting Chief Battlespace Surveillance Innovation Center

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(31-07-2008)*
31-07-2008

2. REPORT TYPE
Final Report

3. DATES COVERED *(From - To)*
29-06-2005 to 30-06-2008

4. TITLE AND SUBTITLE
Improving Magnitude Detection Thresholds Using Multi-Station

Multi-Event, and Multi-Phase Methods

5a. CONTRACT NUMBER
FA8718-05-C-0022

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62601F

6. AUTHOR(S)
D. Schaff and F. Waldhauser

5d. PROJECT NUMBER
1010

5e. TASK NUMBER
SM

5f. WORK UNIT NUMBER
A1

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Columbia University
254 Engineering Terrace
1210 Amsterdam Avenue, MC 2205
New York, NY 10027

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory
29 Randolph Road
Hanscom AFB, MA 01731-3010

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RVBYE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RV-HA-TR-2008-1121

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

We report on the results of both a semi-empirical analysis and a case study in Xiuyan, China, that demonstrated that a correlation detector can lower magnitude detection thresholds by over one full unit for similar events as compared to a standard STA/LTA detector. 90 out of 90 events (100%) were detected by cross correlation for the case study whereas a procedure like the pIDC employs detected only 11%. We found that less than perfect matches still provided useful detections due to location, focal mechanism, and magnitude differences. Events with magnitude differences as large as 2.3 and 3.3 were shown to produce detection spikes on the correlation traces. Further work applied the correlation techniques on a larger scale to 5,000 events at Parkfield, California, and 19,000 events in and near China. We are attempting to see how broadly applicable correlation methods can be applied to different tectonic settings and for what percentage of the seismicity. 111 million correlations were performed on Lg-waves for the events in China at 363 stations. Final results indicate two thirds of the 19,000 events can be detected by cross correlation using this relatively sparse regional network. The full magnitude unit reduction in detection threshold is confirmed on a large scale in China and Parkfield.

15. SUBJECT TERMS
Cross correlation, Detection, Discrimination

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Robert Raistrick
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR	67	**19b. TELEPHONE NUMBER** *(include area code)* 781-377-3726

Table of Contents

Figures

Acknowledgments

We'd like to thank the operators of the Chinese Digital Seismograph Network, the U.S. Geological Survey, and the Berkeley data center for the seismic data used in this project as well as IRIS for the archiving and access to the data. We are grateful to Paul Richards, Felix Waldhauser, Won-Young Kim, Steven Gibbons, Dave Harris, and Bill Menke for helpful discussions.

1. SUMMARY

Statistical analyses were conducted on the capability of correlation detectors for similar events. Semi-empirical synthetic runs took a 50 s window on an Lg-wave recorded at 750 km distance filtered from 1 to 3 Hz and embedded it 300,000 times in real, continuous, background seismic noise. The noise was selected for 36 days spread throughout the year to capture diurnal and seasonal variations. No screening for random, unknown signals in the noise was performed. A correlation detector has a 50% probability of detection with 1.5 false alarms per day for a signal-to-noise ratio (SNR) of 0.32 which corresponds to a full magnitude unit reduction in detection threshold over a standard STA/LTA technique. A scaled cross correlation coefficient performs slightly better with 1 false alarm per day and has fewer false triggers on unknown, random signals. Summing the cross correlation traces together for all three components enhances the detection signal similar to beamforming. A correlation detector summing the correlation traces for the three components together has a 96% probability of detection with zero false alarms in 36 days for a SNR of 0.32. The significant result of this study is that a correlation detector has more than an order of magnitude improvement in detection threshold for similar events with acceptably low false alarm rates to be used in practice. Comparisons are made of the performance of a correlation detector compared to a standard STA/LTA detector for 90 events in the 1999 Xiuyan, China, earthquake sequence. Triggers on three component data verify independent detections to the nearest sample. Unrelated signals that trigger above a threshold do not align on the three components and do not constructively interfere when the cross correlation traces are stacked or averaged. Semi-similar events due to less than perfect matches arising from location and mechanism differences or source complexities can provide useful detections. Large and small events correlate well enough for detection. Two examples are shown--one with a 2.3 magnitude unit difference and one with a 3.3 magnitude unit difference. Aftershocks buried in the coda of mainshocks can be detected. The correlation detector finds 90 out of 90 or 100% of the events whereas a STA/LTA detector finds 10 out of 90 or 11%. This represents a 1.3 magnitude unit reduction in detection threshold for these events. The correlation techniques were then applied on a larger scale to 5,000 events at Parkfield, California, and 19,000 events in and near China. We are attempting to see how broadly applicable correlation methods can be applied to different tectonic settings and for what percentage of the seismicity. 111 million correlations were performed on Lg-waves for the events in China at 363 stations. Final results indicate two thirds of the 19,000 events can be detected by cross correlation using this relatively sparse regional network. For Parkfield 82% of the events studied can be detected by cross correlation. Correlation detection is able to find additional events beyond what standard processing detects for China (70% increase) and for Parkfield (factor of 10 increase like Gutenberg-Richter preidcts). Most event separation distances for events that correlate at Parkfield are less than 1 km. The distribution of magnitude differences for events that correlate at Parkfield is not distinguishable from the input magnitude distribution. Detection magnitude threshold reduction of about 1 unit holds for large scale application to the 19,000 events in China and 5,000 events in Parkfield with false alarm rates of a few percent.

2. INTRODUCTION

This section provides some background and introductory material. In Section 3 we present the methods and procedures used to develop, test, and evaluate the correlation detector techniques and algorithms. Section 4 contains test results and discussion. Concluding remarks are presented in Section 5.

2.1. Project Milestones

Preliminary investigations were made of the multi-event (cross correlation) and multi-station (STA/LTA) methods for improving detection. Based on these results it became apparent that cross correlation could achieve a full magnitude unit in reduction of detection threshold able to identify signals even well below the noise level whereas the multi-station technique appeared to give at most a few tenths of a magnitude unit improvement and was limited to SNRs above unity. Because of the dramatic difference in the performance between the two techniques it was decided to focus the bulk of the research effort on developing, understanding, and testing the correlation method since it was deemed to be the most fruitful return for the labor. This two year project has been broken down into the following four subtasks:

1. Semi-empirical synthetic runs embedding real seismic signals in real background noise to develop technique and evaluate false alarm rates,

2. Case study of ninety events in the 1999 Xiuyan earthquake sequence to observe performance on a real data set,

3. Large-scale application to ~19,000 events in China to quantify reduction in threshold and broad usefulness,

4. Large-scale application to ~5,000 events in Parkfield, California, where we have better control on the locations.

2.2. Background

The existence of similar waveforms for seismic events and the exploitation of cross-correlation techniques has found many wide-spread applications including identification of repeating events, improving hypocentral locations, and detection of lower magnitude events. Much of the history of research in the literature seems to have focused primarily on improving locations (Poupinet *et al*., 1984; Fréchet, 1985; Ito, 1985; Frémont and Malone, 1987; Deichmann and Garcia-Fernandez, 1992; Got *et al*., 1994; Dodge *et al.*, 1995; Nadeau *et al.*, 1995; Shearer, 1997; Lees, 1998; Rubin *et al*., 1999; Waldhauser *et al.*, 1999; Waldhauser and Ellsworth, 2000; Phillips, 2000; Rowe *et al.*, 2002; Schaff *et al.*, 2002; Moriya *et al.*, 2003; Waldhauser *et al.*, 2004; Shearer *et al.*, 2005; Hauksson and Shearer, 2005). Using waveform cross correlation for detection has been studied less. Early work identified and characterized events with master templates of quarry

explosions (Harris, 1991). Subsequent work extended correlators to subspace detectors to allow for variations in waveforms for a given source region (Harris, 1997). Case studies are beginning to show more promise for correlation detectors in reducing magnitude thresholds for practical applications (Gibbons and Ringdal, 2004; Gibbons and Ringdal, 2005; Gibbons and Ringdal, 2006; Gibbons *et al.*, 2007; Schaff and Waldhauser, 2006). The case studies, however, are not able to estimate false alarm rates unless a denser, complete, local catalog is available. Wiechecki-Vergara *et al.* (2001) have derived expected false alarm rates under certain assumptions about the distribution of cross correlation coefficient (CC) values as well as the statistical significance of a given CC. Our work examines false alarm rates and the statistics of detection using empirical distributions of CC values obtained from real seismic signals and background noise.

When monitoring seismic events whether in mining environments or on a global scale, the first order of business is to detect the signals apart from the noise. Without an initial detection no further work of associating events, locating, determining magnitudes, focal mechanisms, etc. can take place. Current operational procedures for many seismic networks employ a power detector where the energy in a short-term average window (STA) is divided by a long-term average window (LTA) and a detection is triggered when this ratio exceeds some signal-to-noise ratio (SNR) threshold (Freiberger 1963). The STA/LTA filter is useful because it works for all signals requiring very little knowledge about its characteristics *a priori* except that the energy exceeds the noise in typically limited, narrow filter bands. The drawback with this method is that false alarm rates go up dramatically for lower SNR thresholds.

At the other end of the spectrum of possible detection techniques is the correlation detector when perfect knowledge of the signal is available (Harris 2006). In the presence of Gaussian, white noise a correlation detector, also known as a matched filter, is the optimal means of detecting a known signal (Van Trees 1968). Such a detector offers amazing sensitivity with the benefit of few false alarms. Gibbons & Ringdal (2006) demonstrated in a case study at the NORSAR array approximately 1.1 magnitude unit reduction in detection threshold using cross correlation compared to an STA/LTA detector. Embedding a known signal in real, background, seismic noise enables an empirical estimate of the false alarms (1 per day, which is acceptably low) that corresponds to this order of magnitude improvement in detection threshold (Schaff , 2008). With so few false alarms this amount of improvement is no small feat. A common SNR threshold for detection using an STA/LTA filter is 3.2 or 10 dB (Ben Kohl personal comm.; http://www.rdss.info/librarybox/idcdocs/pages/521.html). To achieve one magnitude unit reduction means that signals with amplitudes 10 times less for the same amount of noise must be detected which corresponds to an SNR of 0.32. In other words the signal is actually *below* the noise level with an amplitude of approximately a third as great. An STA/LTA filter can only detect events with SNRs above unity and can never extract signals below the noise level. In practice the threshold is much higher than unity to reduce the number of false alarms. For the prototype International Data Center (pIDC) typical thresholds range from 3.0 to 4.5 depending on the station (Ben Kohl private communication; http://www.rdss.info/librarybox/idcdocs/pages/521.html).

The disadvantages of a correlation detector are that similar master waveforms must be available for the region of interest and the computational cost of correlating with many templates for seismically active areas. Both of these problems can be addressed by the

application of subspace detectors (Harris 2006; Harris & Paik 2006). Instead of correlating incoming data streams with single waveforms they are matched with a linear combination of basis waveforms that span the subspace. This allows for less than perfect waveform matches due to location, mechanism, magnitude differences and source-time function complexities. It also reduces the number of templates to a more manageable set. Subspace detectors, in fact, theoretically bridge the gap between the two end-members of the spectrum of possible detectors--STA/LTA filters when all the basis waveforms are used and correlation detectors when only a single basis waveform is compared and receives all the weight (Harris 2006). For our work, however, we examine only how the correlation detector by itself performs and compares to the standard STA/LTA procedure. We also examined how a correlation detector manages for less than perfect matches.

Also worth noting here is the usefulness of cross correlation for other applications such as discrimination.

3. TECHNICAL APPROACH

Figure 1 shows an Lg-wave recorded 750 km away embedded in various levels of real background seismic noise. The *M* 4.3 event occurred on December 1, 1999, 4:45:31.10 (40.6º, 122.7º) in the 1999 Xiuyan earthquake sequence in China. The Lg-wave was chosen because it was found to correlate well for this sequence of events and in the region (Schaff and Richards, 2004a, 2004b). Even though it is an emergent arrival, it is well-suited for a correlation detector because of its large amplitudes, long duration of signal energy, and high frequency content. The station is MDJ with network code IC. The channel is BHZ**.** The waveforms are filtered from 1 to 3 Hz. Signal-to-noise ratios (SNR) are shown to the left and are computed as a mean absolute value like an STA/LTA filter uses. The signals are scaled linearly and no account is made for the corner frequencies of the different magnitude events. A standard trigger used by the Prototype International Data Center (pIDC) of 3.2 or 10 dB (Ben Kohl personal comm.; http://www.rdss.info/librarybox/idcdocs/pages/521.html) shows the signal that a standard detector (STA/LTA) would find in this case corresponding to a magnitude 3.5. The relationship between SNR and magnitude is:

$$\mathrm{Mag} = \log(\mathrm{SNR}) + 4.3 - \log(\mathrm{SNR}_{\mathrm{Mag}=4.3})$$

To reduce the detection threshold by 0.5 magnitude units to 3.0 it can be seen that the SNR drops to one. It is impossible for an STA/LTA filter to detect a signal at the noise level. To reduce the detection threshold by a full magnitude unit to 2.5 corresponds to a SNR of 0.32 or where the signal is one third of the noise level.

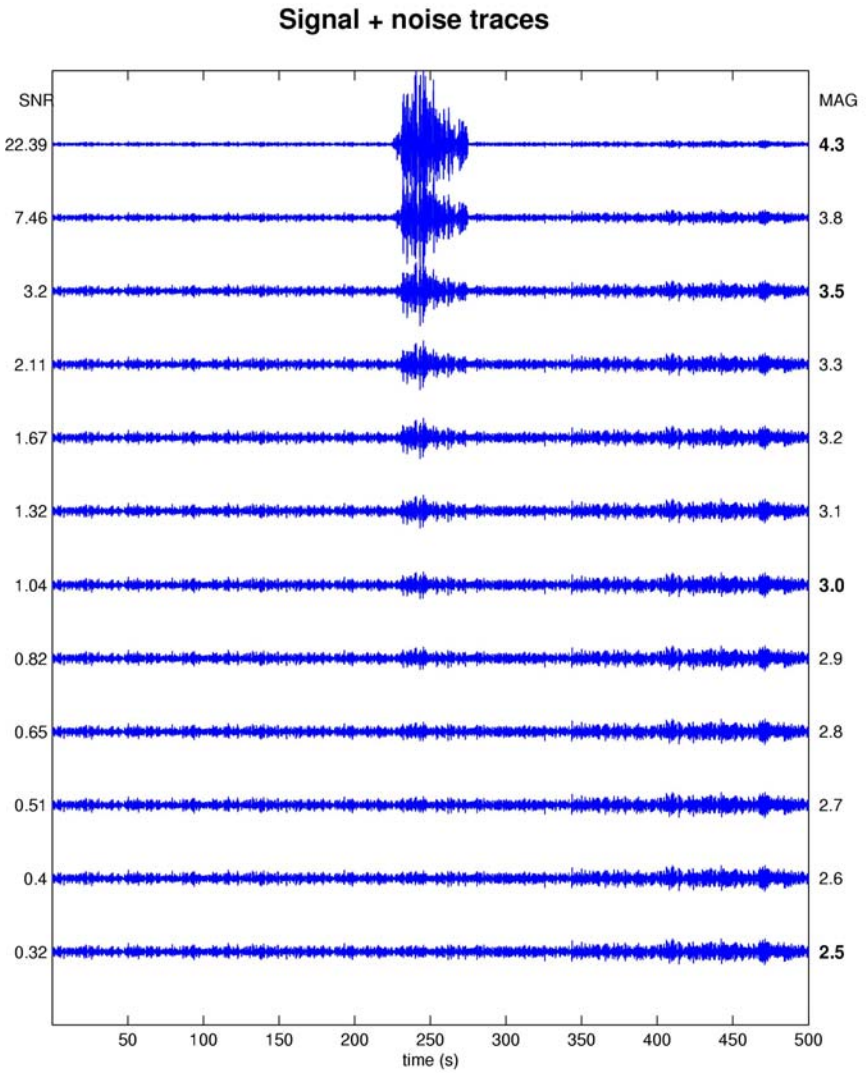

Figure 1. Lg-wave signals for various magnitudes embedded in real, background seismic noise. Signal-to-noise (SNR) values on left show that a normal trigger of 3.2 corresponds to a *M* 3.5. To achieve a full unit reduction in magnitude threshold (*M* 2.5) the signal is no longer visible to the eye and is at one third the noise level (SNR = 0.32).

A magnitude 3.5 event at the same location as a magnitude 2.5 event with similar mechanisms would theoretically have similar seismograms in a pass band below the corner frequency of both events. Gibbons *et al.* (2007) demonstrated that the signal from a M 3.5 master event was able to correlate with and detect events three orders of magnitude lower (confirmed by recordings from local stations). These two statements suggest that our simplifying assumption of linear scaling of the signals is reasonable for our purposes of testing improved detection capability since we are considering only one magnitude unit difference.

Figure 2 displays the cross correlation traces for each of these signals correlated with the 50 s master template, the identical signal without noise. 50 s is chosen because that is the duration of the Lg-wave energy which is primarily dominant in the 1 to 3 Hz

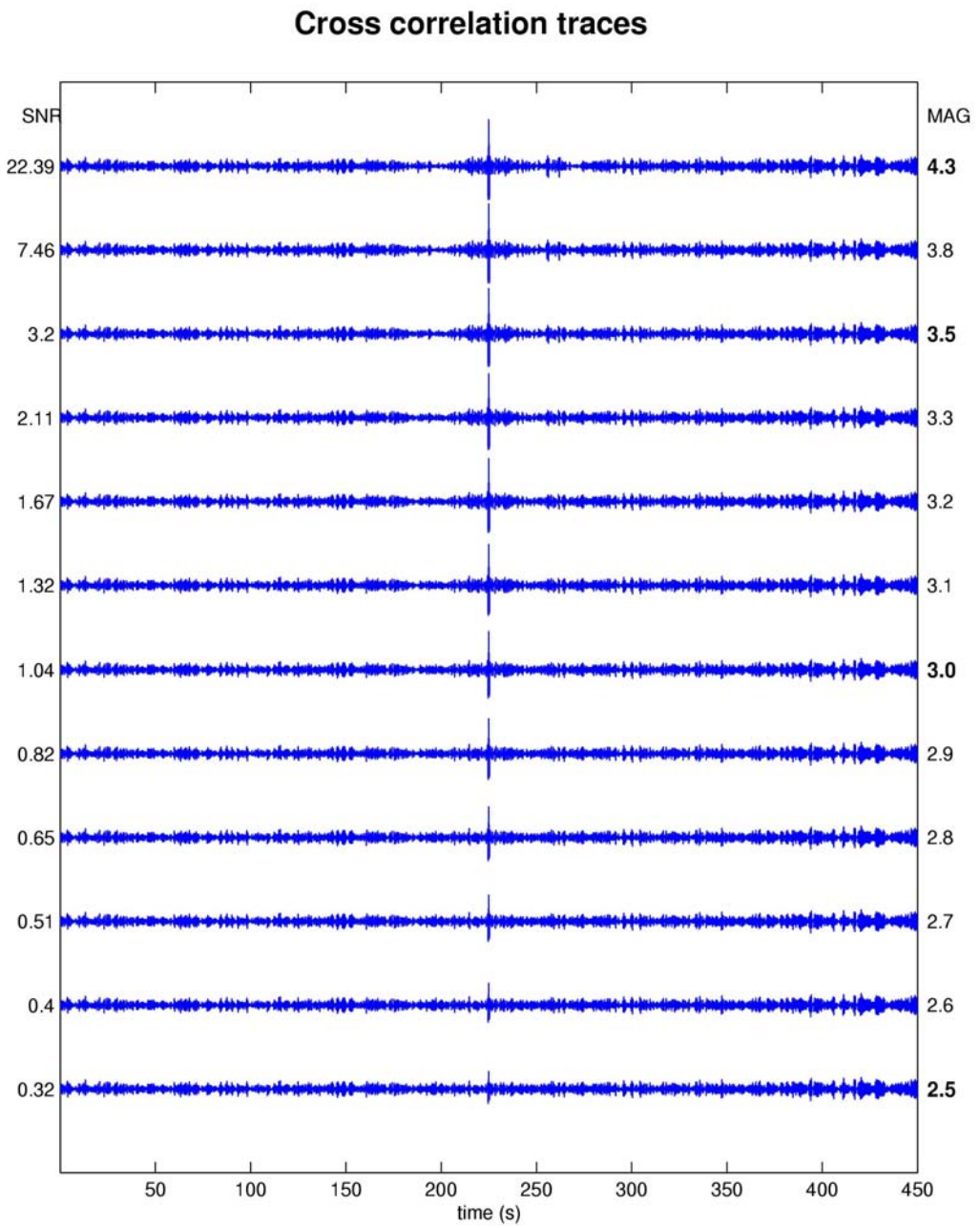

Figure 2. Cross correlation traces with a 50 s master event corresponding to Figure 1. Clear detection spikes are seen at 225 s for all even down to *M* 2.5.

band. Clear detection spikes can be seen all the way down to magnitude 2.5 giving us the first indication that a correlation detector can reduce the detection threshold by a full magnitude unit. In the next section we will explore if such detections occur with acceptably low false alarm rates.

Figure 3 shows the maximum cross correlation coefficient (CC) as a function of SNR highlighting three magnitudes of interest. It can be seen that detecting half a

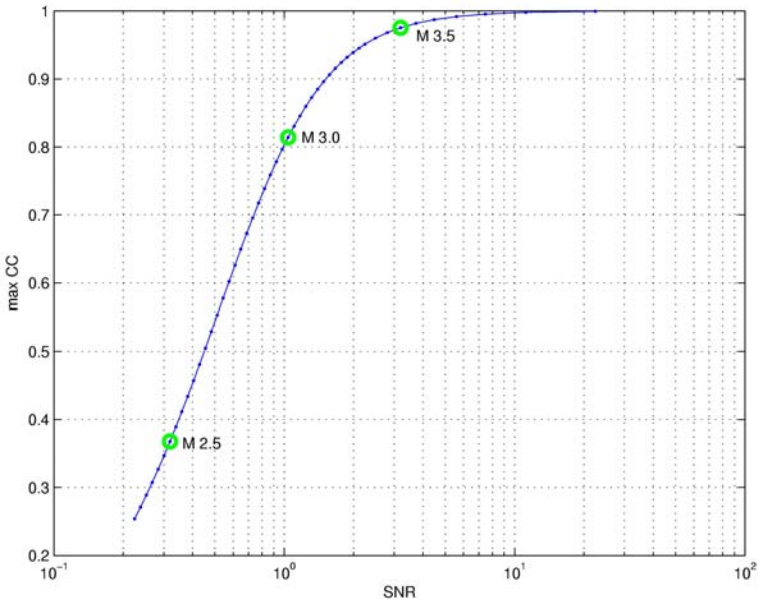

Figure 3. Maximum cross correlation coefficient as a function of SNR from Figure 2. Values for *M* 3.5 and 3.0 would be easily detected. Value for *M* 2.5 is extremely low and would normally be discarded for location purposes.

magnitude unit lower at 3.0 would easily be picked up with a cross correlation coefficient above 0.8. Remember that this corresponds to a signal right at the noise level. To detect a full magnitude unit lower, however, at 2.5 the cross correlation coefficient is quite low at 0.367. Normally this type of measurement would be discarded for purposes of measuring a relative arrival time even though there is a clear detection spike relative to background levels. The reason is explained in Figure 4 with two dissimilar traces in the top panel. The cross correlation function of these two waveforms corresponds to the first trace in the middle panel with a coefficient of 0.364. Therefore even though these two waveforms are unrelated they have a coefficient similar to the case of an identical signal buried in noise (0.367). The second trace in the middle panel is the last correlation trace from Figure 2 corresponding to a SNR of 0.32. A detection threshold of 0.35 would trigger on both of these examples. The first case would be a false alarm, whereas only the second is the true detection that we want to capture. One solution to this problem is based on the observation that in the second case the maximum is high relative to background values. We can apply an STA/LTA filter to the cross correlation traces which is shown in the bottom panel. We choose for the window length of the STA one sample and the window of the LTA 20 s. A delay between the STA and LTA is also employed of 4 samples (0.2 s) to avoid side lobes of the cross correlation trace. Here the maximums differ by a substantial amount (5.8 for the dissimilar case and 8.3 for the identical signal buried in noise). Gibbons and Ringdal (2006) employ a similar procedure which they call a “scaled cross correlation coefficient” (SCC) where they scale the CC by the RMS within some window of the background levels. Because our STA/LTA filter uses mean absolute value we divide by that LTA instead of RMS but the effect is

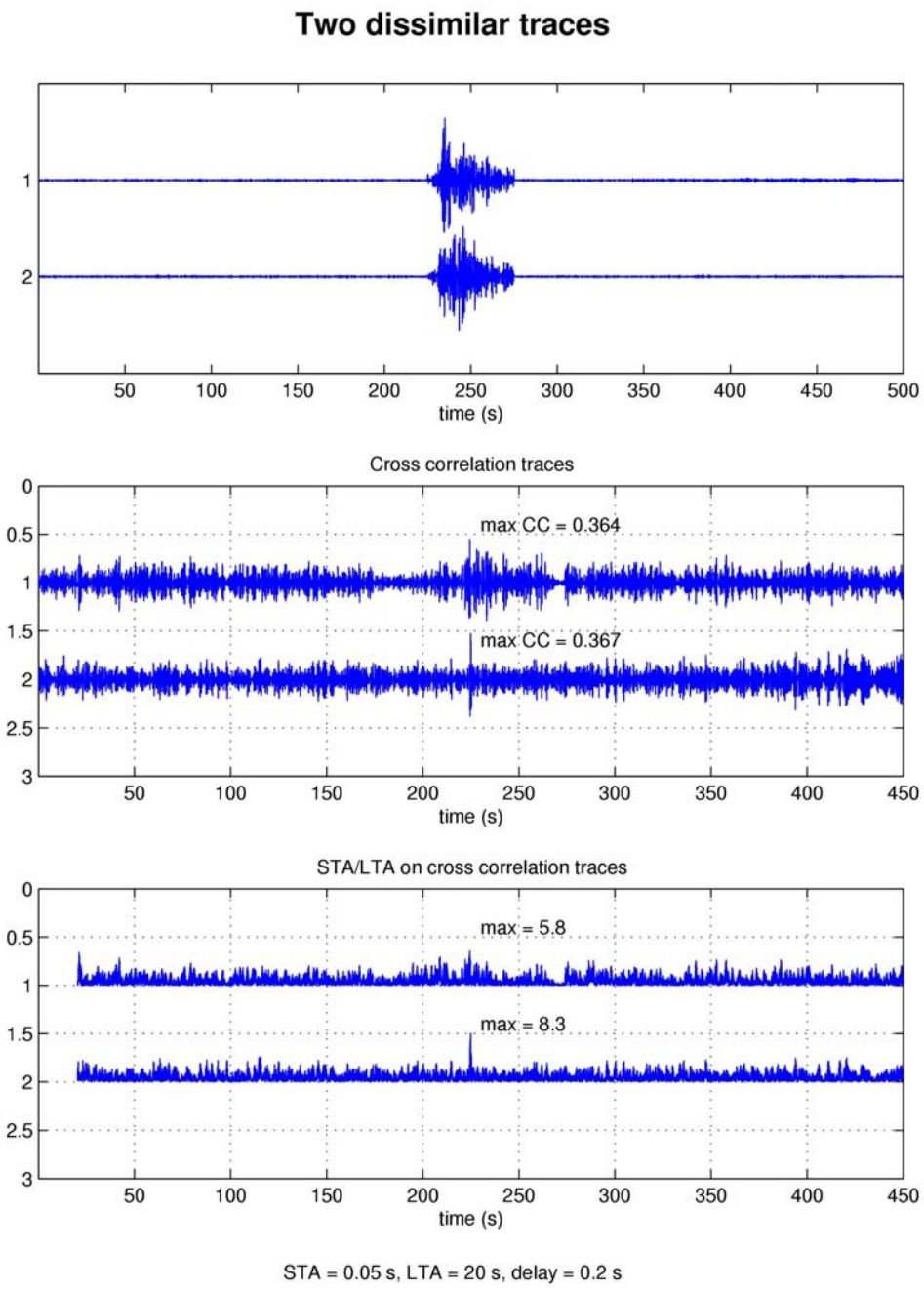

Figure 4. Two dissimilar traces (top panel) would produce a false alarm with CC (first trace of second panel) but not with a scaled CC (first trace of third panel).

basically the same. It can readily be seen that a scaled CC threshold of 6 would weed out the dissimilar case and would leave us with the true detection.

Next we examine how the correlations perform for all three components. In Figure 5 we cross correlate signals for the BHN, BHE, and BHZ components in the top panel with the same signals plus noise in the second panel (SNR = 0.32) to obtain the

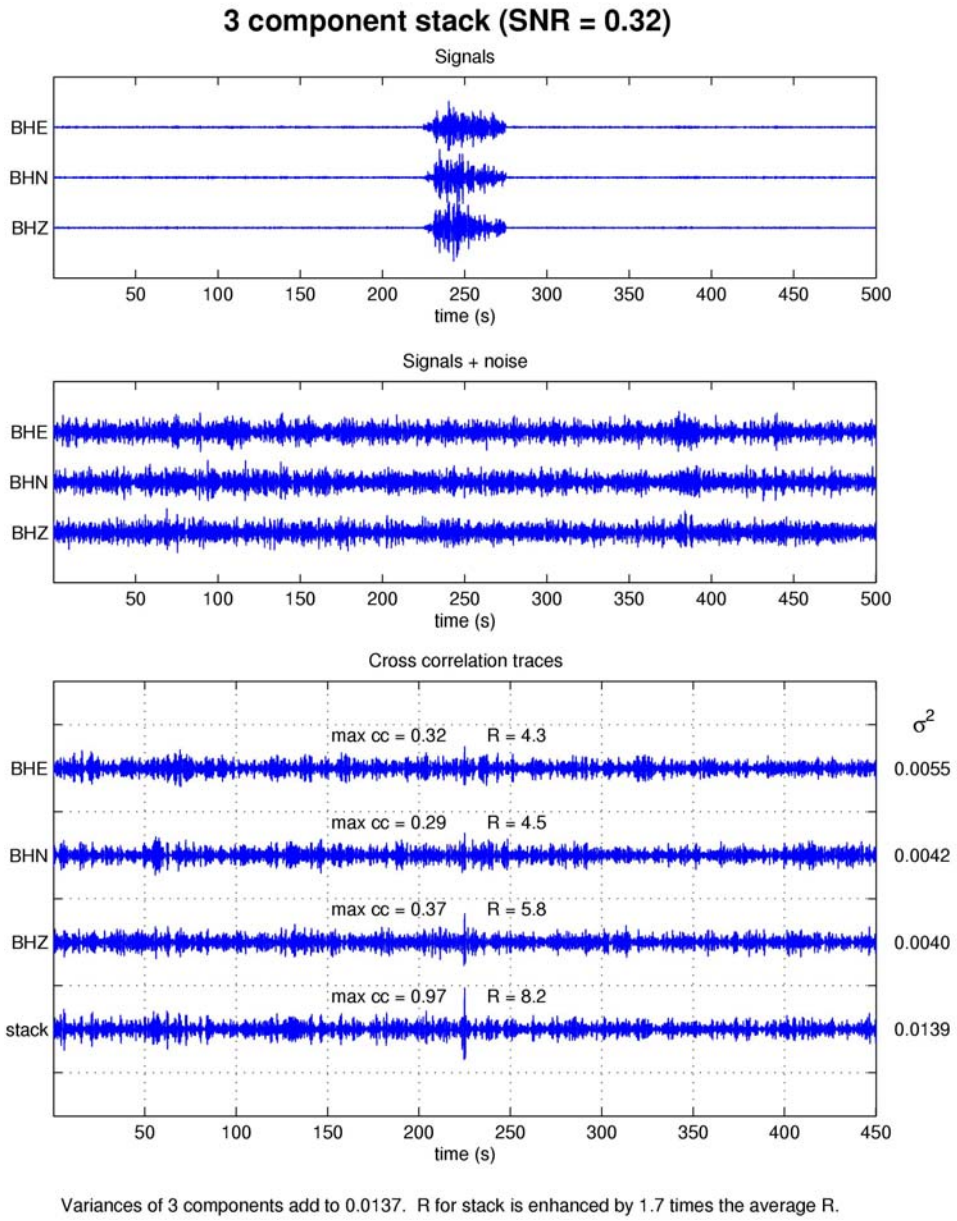

Figure 5. Averaging three component cross correlation traces for BHE, BHN, and BHZ channels enhance the detection spike similar to beamforming on arrays.

cross correlation traces in the third panel with CCs around 0.3 as before and clear detection spikes at 225 s. If we average the cross correlation traces we see that the spikes constructively interfere and give a new maximum of 0.32. There is an overall enhancement in the spike because the background levels deconstructively interfere. We can estimate the level of this enhancement. Define,

$$R = \frac{\max cc}{\sigma}$$

Assume σ^2 is the variance for all three components. For the stack the variances add, $\sigma^2_{stack} = 3\sigma^2$. Then,

$$R_{stack} = \frac{3\max cc}{\sqrt{3\sigma^2}} = \sqrt{3}R = 1.732R$$

We see that the variances of the three components do add to 0.0137 which is close to 0.0139 of the stack indicating that they are approximately normally distributed. Also R_{stack} is enhanced by 1.7 times the average R as expected. (Note: that max cc for the stack is 0.97 for calculating R. This doesn't imply near perfect semblance, but is just a coincidence since they are summed, the average is 0.32.) This is the similar to the improvement in signal enhancement that is achieved by beamforming on arrays, stacking the seismograms themselves.

There are, however, subtle differences between stacking the seismograms and correlation traces. Stacking seismograms to increase SNR is only useful when the signals constructively interfere which is not always the case as inter-station distance increases. Signal enhancement averaging cross correlation traces is always without loss regardless of array geometry (Gibbons and Ringdal, 2006) and/or component (i.e. orientation of particle motion). Therefore stacking the seismograms themselves on three components is pointless since the particle motions are different.

3.1. Independence of Scaled CC

It is known that CC depends on both window length and bandwidth. In the top panel of Figure 6 CC is seen to increase dramatically as window length decreases over three orders of magnitude. The curve is constructed from the average of one thousand correlations of random noise traces for each window length. If the windows and bandwidth were infinite then CC should be equal to zero which is the asymptote of the curve for increasing window length. It can be seen for a CC of 0.35, such as we have examined above for a threshold, a window length of 50 samples would on average trigger a detection for purely noise traces. CCs as high as 0.7 are seen for window lengths of 10 samples. This would produce many false alarms for typical detection thresholds. Scaled CC on the other hand in the lower panel of Figure 6 is relatively flat over this wide range of window lengths. It in fact shows a slight decrease from 100 sample windows to 10 sample windows. The scale of the y-axis is from 0 to 6 (the empirical threshold we have determined for scaled CC). Therefore this behavior demonstrates that there is little chance of scaled CC producing false alarms on average for purely noise traces as a function of window length.

CC also depends on the filtering of the waveforms. Figure 7 displays CC as a function of bandwidth in the top panel. At the right for a bandwidth of 9 Hz the purely noise waveforms are filtered from 1 to 10 Hz for a window length of 500 samples. Then as bandwidth decreases to zero, the low and high corners are shrunk on both sides towards 5.5 Hz. CC approaches unity for the case of the single frequency 5.5 Hz. This is because the narrow band-pass filter at a single frequency is just a sine wave and two sine waves correlate perfectly differing only in phase. If the bandwidth is 1 Hz the average CC is 0.5 which would trigger many false alarms on random noise traces. One should note that although higher frequencies and larger bandwidths are desirable for CC thresholds, the seismic signal may only contain energy in a limited band. Also the more

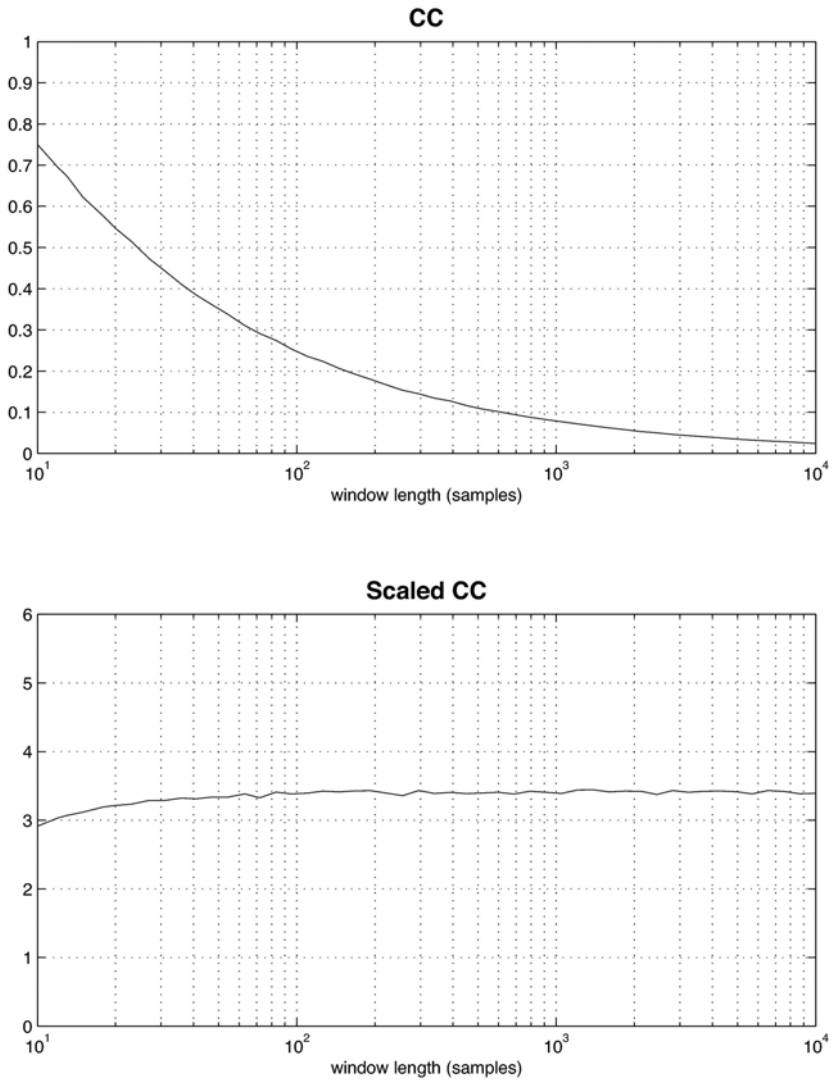

Figure 6. Strong dependence of CC and relative independence of scaled CC on window length.

high frequencies are used the smaller the spatial scale events will correlate over and the less successful the master event will be able to detect over broader regions. Scaled CC in the lower panel of Figure 7 shows a slight decrease with decreasing bandwidth from 3 to 2. Again the opposite trend of this dependence is beneficial for producing fewer false alarms. And the maximum values are significantly less than the threshold of 6 for scaled CC that we are using. Parameters used for the scaled CC for both the window length and bandwidth comparisons are a STA of one sample, LTA of 20 sec on 20 Hz data, with a delay of 4 samples between the STA and LTA.

It is important to realize that a large time-bandwidth product for the signals is still desirable and ascribes greater statistical significance to a large CC or scaled CC value. The point with the relative independence of scaled CC on processing parameters is that

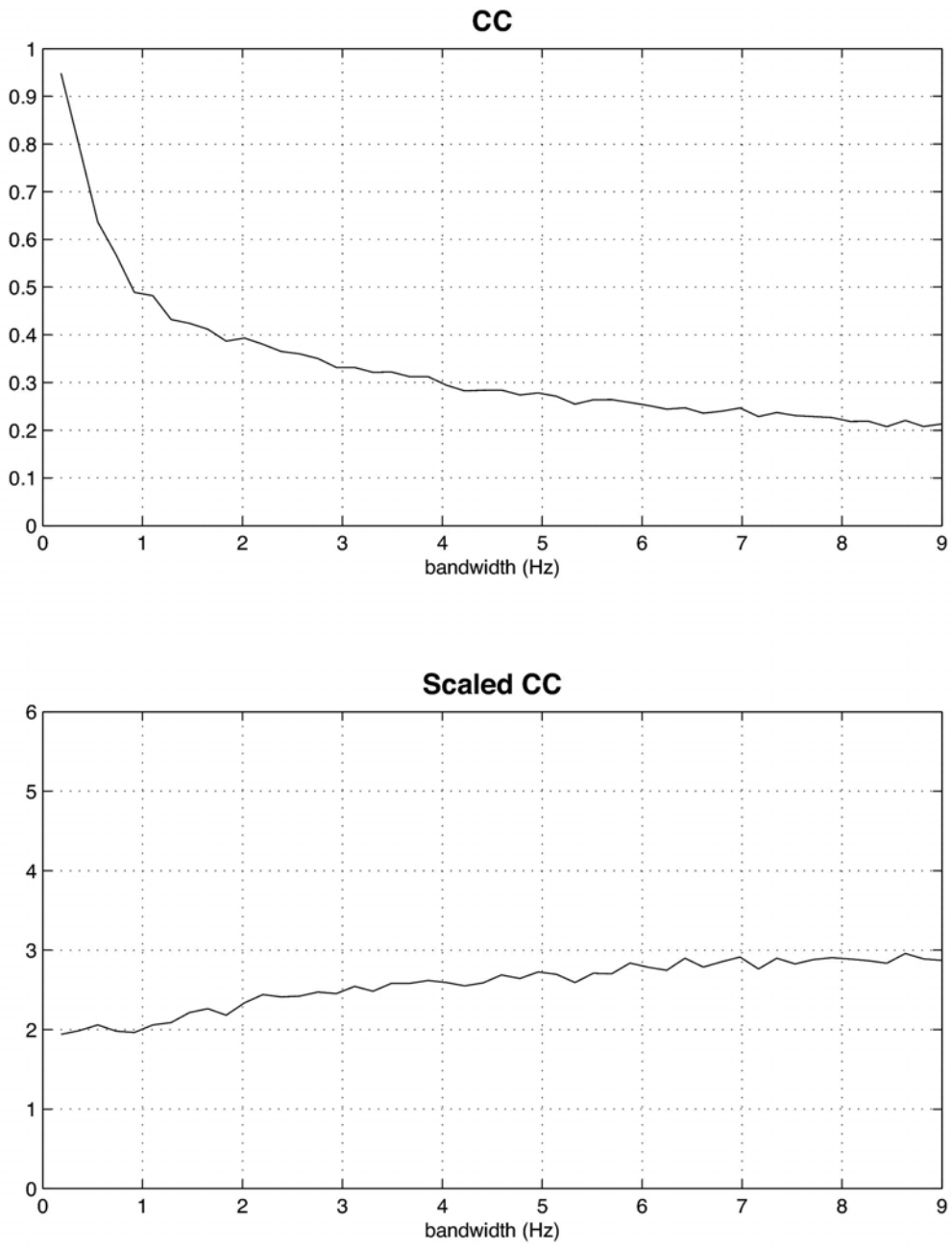

Figure 7. Similar strong dependence of CC and relative independence of scaled CC on bandwidth.

random or noise segments will be less likely to produce false alarms for a given threshold.

4. RESULTS AND DISCUSSION

The results of our project were demonstrated during a series of program and Monitoring Research reviews. These results are summarized below. Material of more interest to the general geophysical community was presented at the annual SSA meetings. (Schafff and Waldhauser, 2006; Schaff, 2007). One paper has been published on the semi-empirical synthetic runs (Schaff, 2008). Another paper is in review at Geophysical Journal International on the case study for Xiuyan. Two other papers are in preparation, one for the large-scale application to China and the other for Parkfield, California.

4.1. Semi-Empirical Synthetic Runs

To look more in detail at statistics of detection and false alarm rates, we take 36 days of actual seismic noise at station MDJ spread throughout the year of 2002 to capture diurnal and seasonal variations. The noise also contains random seismic signals of unknown origin that have not been removed. For the master signal we choose a 50 s window on the Lg-wave of Figure 1 recorded at 750 km distance. The waveforms are all filtered from 1 to 3 Hz. The sample rate is 20 Hz. We embed the signal in the noise ~300,000 times at different intervals in the time series. The noise comprises 62 million samples. We need such large numbers for the statistics because for the false alarm rates we want of about one per day the probabilities are on the order of one in ten million. This is necessary to examine the details of the tails of the distributions. Figure 8a shows the histograms for the signal and noise distributions for CC for a SNR of 0.32. There is a clear separation between the two allowing for detections to be made with a certain threshold. The mean CC is about 0.35 for the signal buried in noise as before. It is instructive to plot probability of detection as a function of probability of false alarm to show how the two trade off in the lower panel which can be computed from the top panel. For a probability of detection of 0.5 there is less than one in a million chance of a false alarm. Given the number of samples per day at 20 Hz this corresponds to 1.5 false alarms per day (Table 1). This is a reasonable false alarm rate and so we therefore conclude that a correlation detector is able to detect a signal buried in the noise one full magnitude unit lower than a standard detector. Figure 8b shows the same signal and noise distributions for a SNR of 0.32 but this time for the scaled CC or signed STA/LTA on the CC trace. Again a clear separation of the distributions is obvious. The mean of the signal buried in noise is around 6. This time, however, the probability of detection at 0.5 corresponds to a slightly lower false alarm rate of one per day.

Scaled CC performs better than CC in terms of false alarm rates for probabilities of detection of 0.63 and lower which corresponds to thresholds of CC = 0.32 and SCC = 6.2. Note here we are preserving the sign of the CC for the SCC so that false triggers don't occur for large negative correlations. For a probability of detection of 0.2 CC has 0.4 false alarms per day whereas SCC has 1/9 or 0.11 false alarms per day (Table 1). However, for a probability of detection of 0.8 CC has 10 false alarms per day while SCC

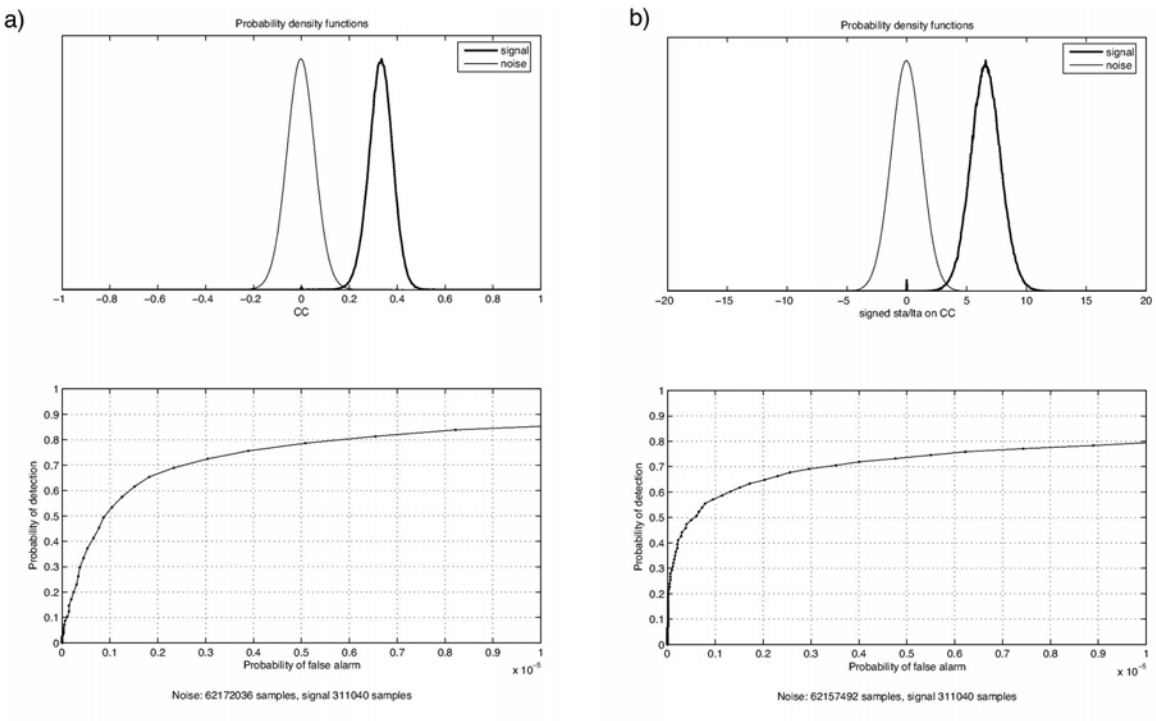

Figure 8. a) Statistics for a signal buried in noise (SNR = 0.32) for CC. Top panel shows a clear separation of the distribution values for known signals from those where noise is present. For a probability of detection (lower panel) of 0.5 there are correspondingly about 1.5 false alarms per day. b) Statistics for a signal buried in noise (SNR = 0.32) for scaled CC. A probability of detection of 0.5 is now 1 false alarm per day.

is eighteen times greater at an unacceptable 184 false alarms per day (Table 1). The reason for this is uncertain. What this means in plain language is that to detect the smallest possible events with the least amount of false alarm triggers on random unknown signals scaled CC can be used. If instead the greatest probability of detecting the most signals of a given magnitude is desired, then the standard CC is better.

If we increase the SNR to 1.012 corresponding to half a magnitude unit reduction in detection threshold over an STA/LTA filter then we see in Figure 9a that the signal and noise distributions are extremely well separated. The mean CC for the signal is above 0.7 indicating a high degree of similarity. In this case there are zero false alarms for the entire 36 days considered at a probability of detection of 99.996%. It is remarkable that even though the signal is at the noise level there is nearly a 100% probability of detection with a zero false alarm rate. For scaled CC the probability of detection is slightly lower (99.8%) at the same false alarm rate consistent with the results before (Table 1).

Figure 9b is the same as Figure 8a for a signal buried in noise at the SNR = 0.32 level except all three components are used to enhance the detection as shown in Figure 5. Note that the mean CC is about the same but that the width of both the signal and noise distributions has been reduced by summing the variances. The narrower distributions cause less overlap of the probability density functions. This time there are zero false alarms in 36 days with a probability of detection of 96.5% compared to the one false alarm per day rate at the 50% level before. If instead we average the SCC traces for the

Table 1. Correlation detector statistics

SNR1	SNR2	Mag diff	$Prob_{detection}$	$Prob_{false}$	False rate	Type	Thresh
Inf	0.32	1	0.5	9e-7	1.5/day	CC	0.34
Inf	0.32	1	0.2	2e-7	0.4/day	CC	0.38
Inf	0.32	1	0.8	5.8e-6	10/day	CC	0.29
Inf	0.32	1	0.5	6.1e-7	1/day	SCC	6.6
Inf	0.32	1	0.2	6.4e-7	1/9 days	SCC	7.7
Inf	0.32	1	0.8	1.065e-4	184/day	SCC	5.6
Inf	1.012	0.5	0.99996	0	0/36 days	CC	0.42
Inf	1.012	0.5	0.998	0	0/36 days	SCC	8.7
Inf *	0.32	1	0.965	0	0/36 days	CC	0.27
Inf *	0.32	1	0.996	0	0/36 days	SCC	3.9
Inf *	0.16	1.3	0.5	7.3e-7	1.25/day	SCC	3.45
Inf *†	0.16	1.3	0.5	2.65e-6	4.6/day	SCC	6
1.012 *	0.32	0.5	0.83	6e-7	1/day	CC	0.2

* Three component enhancement
† Three component CC traces are first averaged and then scaled

three components we have a probability of detection of 99.6% with zero false alarms (Table 1).

The question now arises, taking advantage of three component enhancement, what is the greatest level of magnitude unit reduction that can be achieved with an acceptably low false alarm rate. We found that reducing the signal amplitude by another factor of two (in addition to the ten before) to 0.16 produces a false alarm rate of 1.25 per day for scaled CC averaged over three components (Table 1). This corresponds to a possible 1.3 unit reduction in magnitude threshold. Interestingly if the CC traces are first averaged over the three components and then scaled, this performs slightly worse with 4.6 false alarms per day (Table 1).

A more realistic scenario considers that the master trace has seismic noise too in addition to the candidate event. Figure 9c displays the same set of curves for a master event that has a SNR of 1.012 and the candidate event that has a SNR of 0.32. The three components are summed together as in Figure 9b. Remarkably even though the master signal is at the noise level and the candidate event signal is one third of the noise level, there is an 83% probability of detection with one false alarm per day. This scenario is more of an extreme case. Realistic SNRs for the master event are probably more representative of the case in Figure 9b.

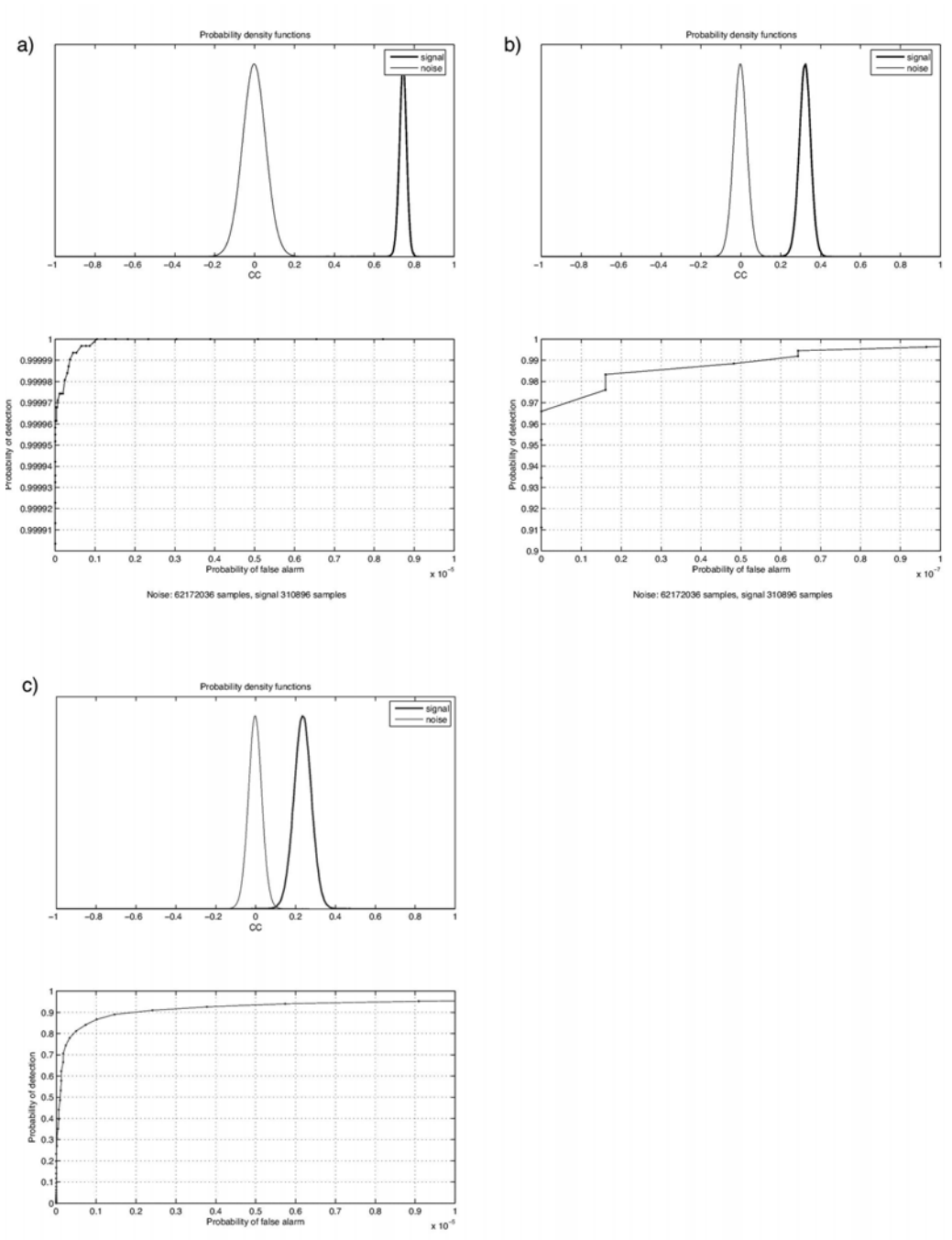

Figure 9. a) Statistics for a signal buried in noise (SNR = 1.012) for CC show wide separation of histograms and almost 100% probability of detection with no false alarms for half a magnitude unit reduction in threshold. b) Statistics for a signal buried in noise (SNR = 0.32) for CC with three component enhancement now achieving 96.5% probability of detection with zero false alarms for a full magnitude unit reduction in detection threshold. c) Statistics for a master event (SNR = 1.012) and a candidate signal buried in noise (SNR = 0.32) for CC with three component enhancement.

4.2. 1999 Xiuyan Case Study

A case study of ninety events is carried out for the 1999 Xiuyan, China, earthquake sequence recorded at stations 500 to 1500 km away (Figure 10). The events studied come from the Annual Bulletin of Chinese Earthquakes (ABCE) which is derived from a much denser network of stations than those that we have waveforms for which are archived at IRIS. We follow this procedure for two reasons. By using a catalog of known events on a denser network we can test what the percentage of events detected by both a correlation detector and an STA/LTA filter is on the sparser network. This is especially important for nuclear monitoring purposes where it is of primary importance not to miss any detections. Secondly, by having independent magnitude estimates to a sufficient level of completeness we can quantify empirically what the reduction in magnitude detection threshold is between the two techniques.

The cross correlation matrix for station IC.BJT is shown in Figure 11. The windows chosen are 50 s long and centered on the Lg-waves filtered from 0.5 to 5 Hz. Clusters of similar events appear with warm colors as blocks on the diagonal. Several values of the CC are quite high above 0.8. Other colors in cyan are at the 0.35 range that are at the level for a small signal buried in noise (SNR = 0.32, Schaff, 2008). The question is whether these values provide reliable detections. Figure 12 shows the cross correlation traces for the events in the cluster from indices 57 to 79, as compared to event number 56. It is seen that there are clear detection spikes on the vertical, north, and east components. In addition, when we look at the average of the three components it can be seen that the spikes constructively interfere and are enhanced compared to the background levels. This is the clearest indication that we have a true detection for these events – the fact that the spikes all align to the nearest sample for basically three independent tests, one for each component. Similar behavior for the other large cluster is shown in Figure 13.

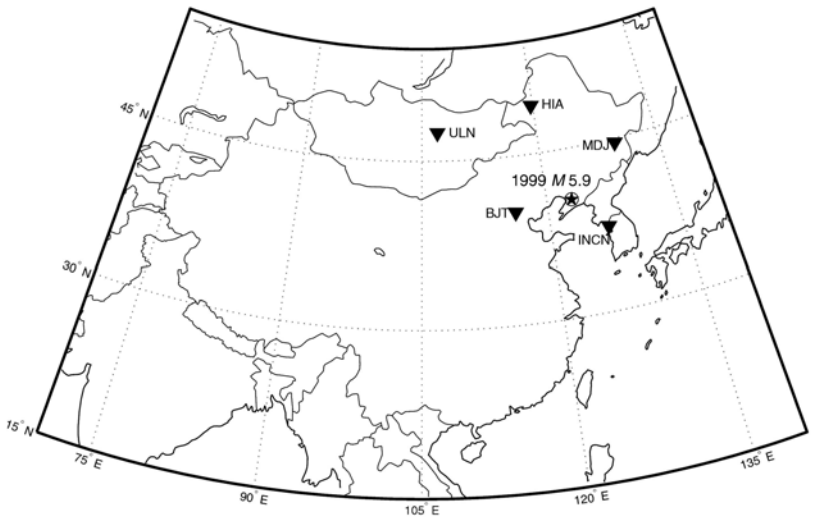

Figure 10. Location of a 1999 Xiuyan earthquake sequence (star). Openly available regional stations (archived at IRIS) recording the events are from 500 to 1500 km away (triangles).

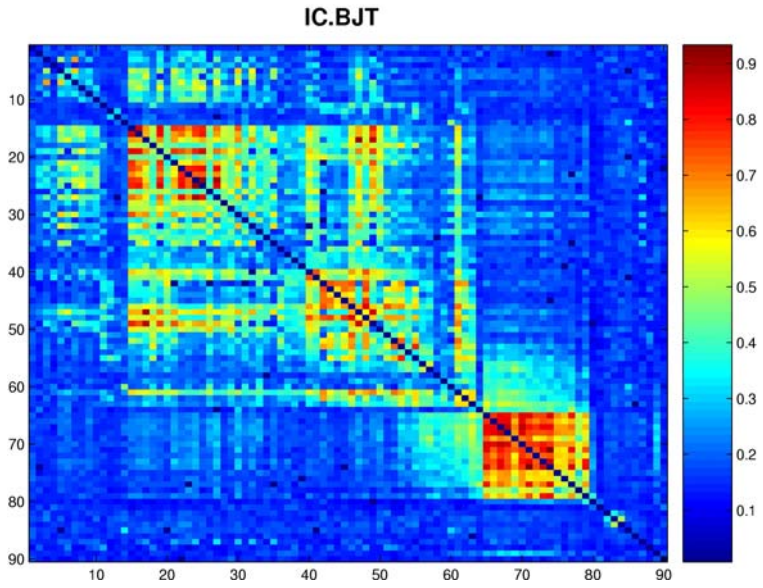

Figure 11. Cross correlation matrix for the 90 events at IC.BJT. Event index is on each axis.

An unrelated event does not show this behavior on all three components. The top panel of Figure 14 shows the waveforms for a non-Xiuyan regional signal with significant amplitudes. The bottom panel shows the correlation traces compared with a Xiuyan master event for all three components in different colors with the maximum coefficients annotated at their time of occurrence. We see that even though one of the individual correlation traces has a maximum at 0.34 which is close to the trigger level for an SNR = 0.32, the maximums on the three components are hundreds of seconds apart and do not align to the nearest sample. When the correlation traces are averaged the result (0.16) is much smaller than the threshold for CC and so no false alarm would be triggered which is how we want the correlation detector to perform for dissimilar events.

To automatically detect the spikes above the background levels we use a scaled CC (SCC) using a threshold of 6 which quantifies the deviation of the cross correlation coefficient from an empirical distribution of background values based on a moving window throughout the correlation trace (Schaff, 2008). Each point in the cross correlation trace is scaled by the mean absolute value of the moving window before the point. Another advantage of using SCC is that it is less dependent on the frequency band and window length than CC (Schaff, 2008). A threshold of 6 corresponds to approximately one false alarm per day on continuous data or a probability of a false alarm being less than one in a million (Schaff, 2008). Pg and Pn phases had 15 s window lengths and 10 s lags that were searched over forwards and backwards. The P-waveforms were filtered from .75 to 2 Hz. Lg phases had 50 s windows with 30 s lags and were filtered from 0.5 to 5 Hz.

Running the correlation detector for all five regional stations and the phases Pg, Pn, and Lg gives the results shown in Figure 15. A blue dot means that the event pair in the matrix satisfied the detection threshold of 6. Again similar events are arranged to be blocks on the diagonal. Note that each event is used as a master event in the matrix to

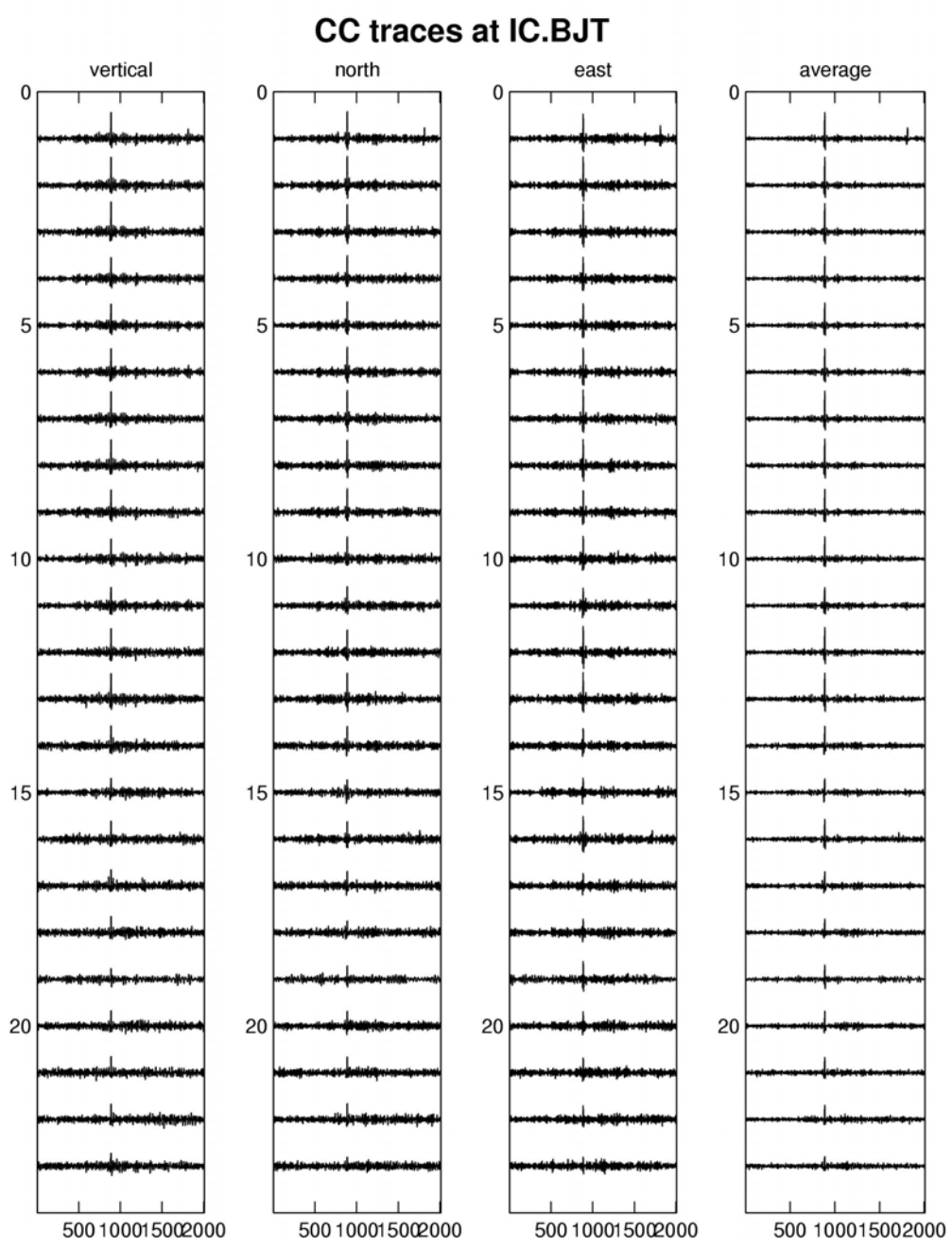

Figure 12. Cross correlation traces for events 57:79 in Figure 11. Master event is 56. Vertical, north, east, and average of three components is shown.

find the maximum possible number of detections. The number beside each phase is the number of events at that station matching the criterion out of 90. The largest amplitude Lg-wave that also has the longest duration of energy in the window produces the best detections and at station MDJ detects 90 out of 90 or 100% of the events. Stations BJT and HIA also have a high number of detected events for Lg. Pg is the next most detected phase and then Pn. Multi-phase detections and detections at multiple stations further ensure that these are true detections. Note the absence of detections for the P-phases is

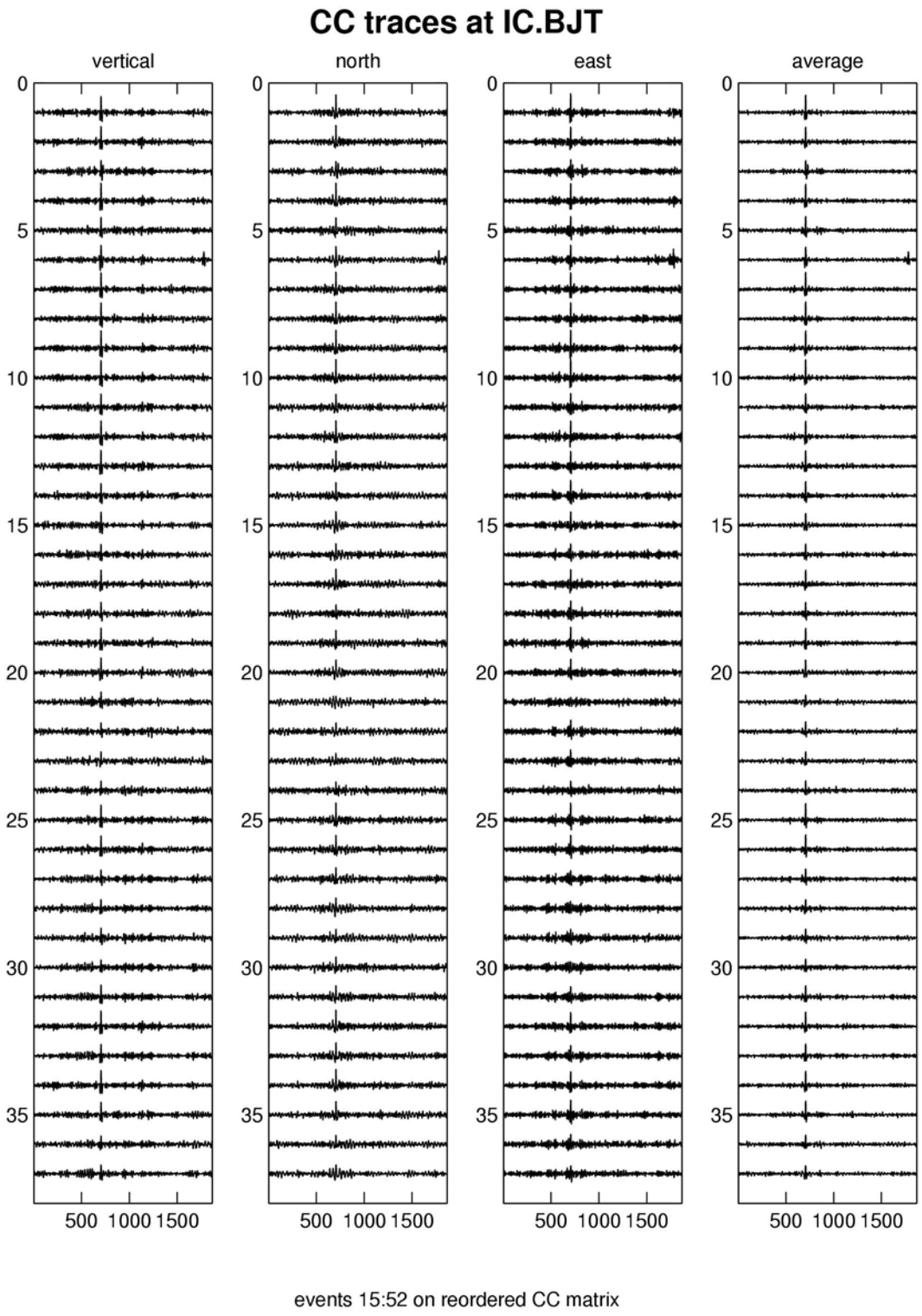

Figure 13. Cross correlation traces for cluster of events 16:52. Master event is 15.

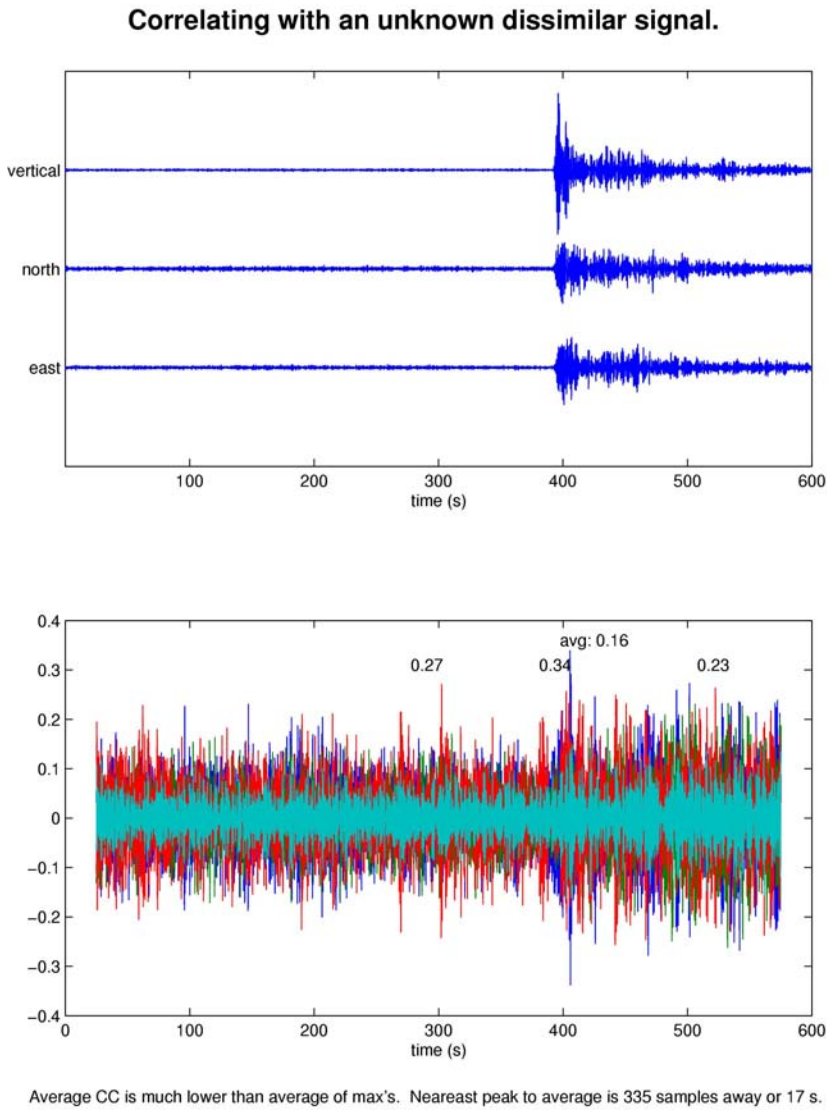

Figure 14. (top) Waveforms for an unrelated signal on three components. (bottom) Correlation traces for the three components with maximum coefficients annotated. Average of three traces in cyan destructively interferes and has a lower maximum (0.16) than on the individual traces.

also informative since it indicates there aren't lots of false alarms triggering on everything and fillling up the matrix.

The top panel of Figure 16 shows a histogram of the magnitudes of the events detected by a STA/LTA filter with procedures like that employed by the PIDC. A STA window of 1 s and LTA of 30 s are used. Trigger levels at the PIDC range from 3.0 to 4.5 depending on the station. We choose 3.2 or 10 dB as a common level. A detection is made only on P-waves on the vertical component. If the trigger is within 10 s of the expected P-wave arrival it is counted as a detection. Overlapping filter bands of 0.5-1, 0.75-1.5, 1-2, 1.5-3, 2.5-5, 4-8 Hz are searched. Finally we used the criterion that 3 or more stations are triggered to form an event detection. With these criteria 10 out of 90 or

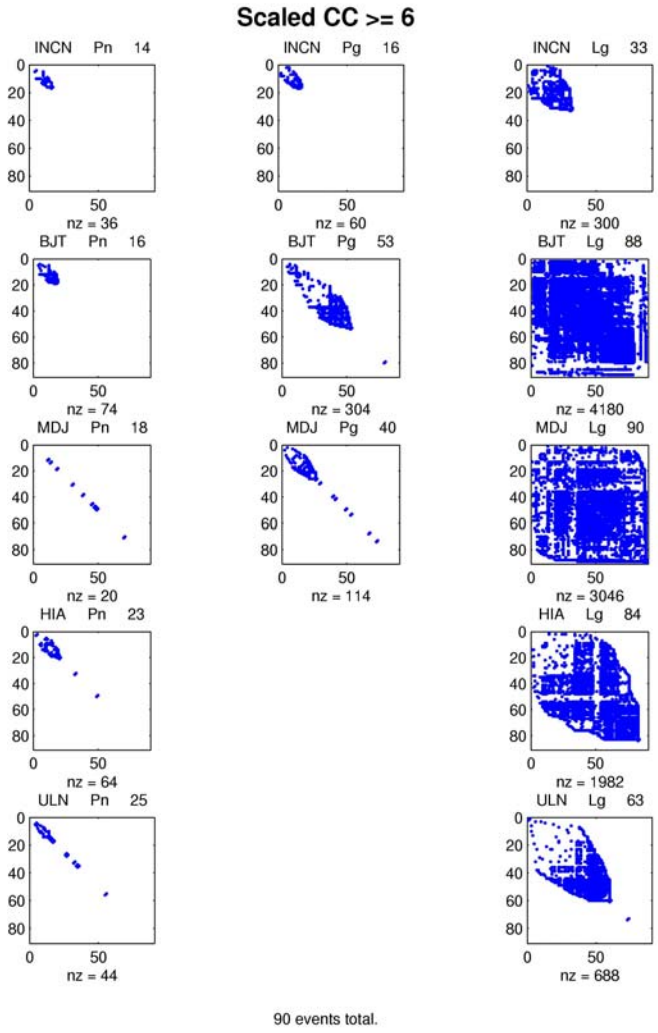

Figure 15. Detection matrices for a scaled CC >= 6 for Pn, Pg, and Lg at five stations. Note largest amplitude Lg gets most detections with 90 out of 90 events at MDJ or 100%.

11% of the events at Xiuyan are detected. Five out of ten of them had magnitudes listed in the Annual Bulletin of Chinese Earthquakes (ABCE) with the minimum being 4.3. The bottom panel of Figure 16 shows the distribution of magnitudes for the events that were detected by cross correlation in Figure 15. Fifty out of the 90 had available magnitudes with the minimum being 3. This represents a 1.3 reduction in magnitude threshold for these events. It can be noted that the ABCE bulletin is not complete below magnitude 3 so some of these events may have actually had lower magnitudes and therefore the reduction in threshold could be even greater.

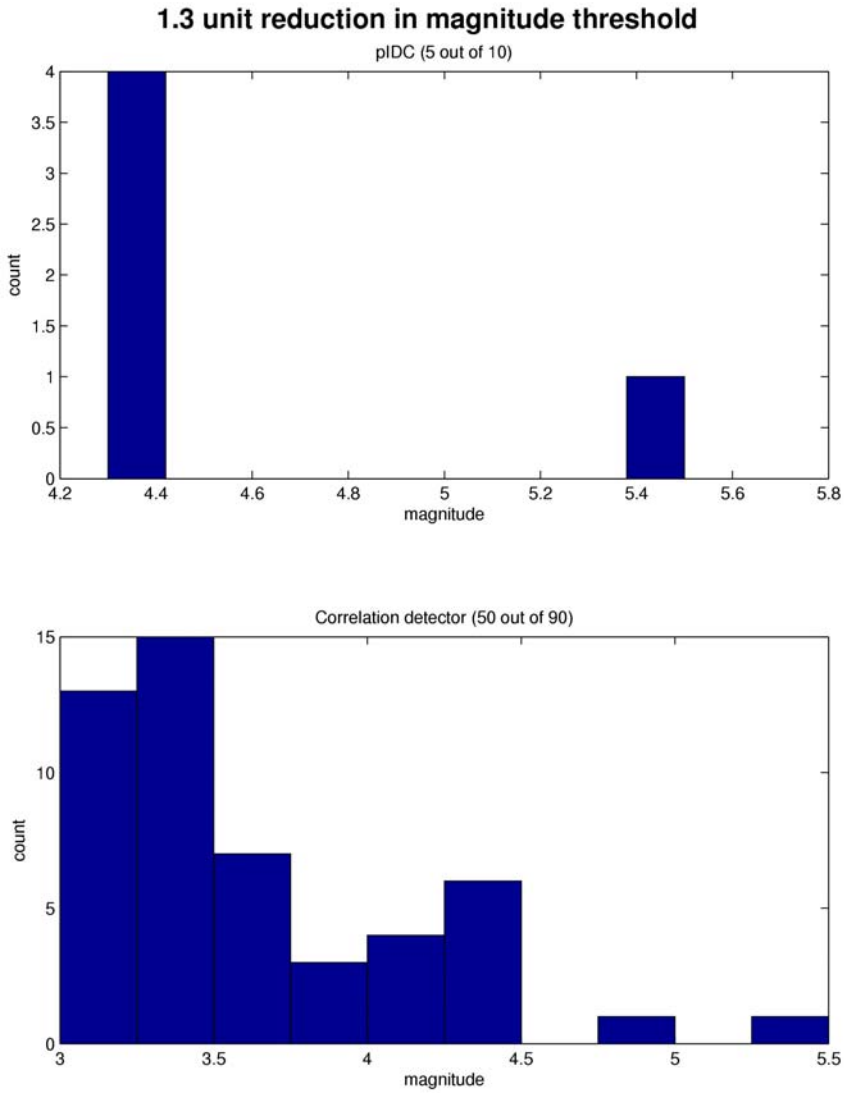

Figure 16. Comparison of STA/LTA detector like that at PIDC with a correlation detector. Magnitude detection thresholds are 4.3 and 3, respectively or a 1.3 unit difference.

4.2.1. Magnitude Dependence of CC

Events can have a low cross correlation coefficient for one of two reasons. First the underlying waveforms can be dissimilar yielding a low value. Secondly the underlying waveforms can be identical but one is buried in significant noise. If the latter is the case the relative magnitude of the slave event can be determined from an amplitude scaling factor, α, (Gibbons & Ringdal 2006):

$$\alpha = \frac{x \cdot y}{x \cdot x}$$

where x and y are the vectors of data for the master and slave events, respectively. We note that this equation is identical to the unnormalized cross correlation coefficient divided by the inner product of the master waveform. The relative magnitudes are then given by the logarithm of the amplitude factor. We estimate the magnitudes of events 56 through 79 from Figure 11 using this procedure fixing the absolute magnitude of the master event and compare to the magnitudes in the local catalog for these events. Figure 17a shows that the two estimates agree with each other. The y=x line is shown for comparison. If we rearrange the order of these 24 events by decreasing average CC we see a gradual decrease in estimated magnitude (Figure 17b) consistent with our assumption that the CC dependence is a function of magnitude and not breakdown in waveform similarity. With the exception of event 16 (M 5.8) which upon examination shows that its large amplitude is real and the lower CC values are likely due to other factors such as source complexities. Figure 17c shows what the CC matrix looks like for this cluster of events. Note how the colors change from red to blue going down the diagonal and from red to cyan moving from the top left hand corner along the axes away from the diagonal. This pattern is strikingly different from what is seen for clusters of events with slight location or mechanism differences. In that case, blocks of warm, red colors corresponding to high CC values appear clustered on the diagonal (pairs within a cluster) and off-diagonal elements are cyan (inter-cluster pairs) which are less similar. In the former case with blue on the diagonal small magnitudes with a greater amount of relative noise correlate with other small magnitudes also with more noise and so yield the lowest

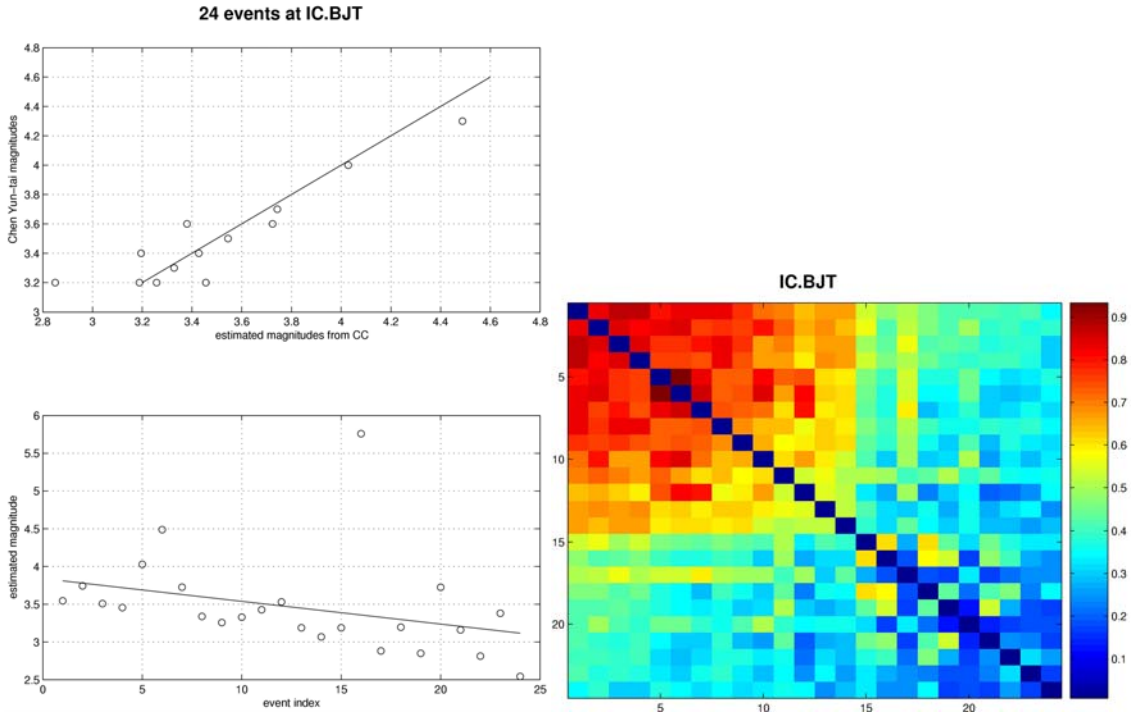

Figure 17. a) Comparison of magnitudes from local catalog and those estimated from unnormalized cross correlation coefficient. b) Events reordered with decreasing average CC show a gradual decrease in magnitude. c) Pattern of a CC matrix for a set of events that show a magnitude dependence to the CC.

CC values. From a single pair-wise CC a low value can't be differentiated whether it is due to location or mechanism differences or if it is a magnitude dependence to the CC. However, looking at the pattern of CC values for a set of events more information is available which can help distinguish between the two cases whether the variations are due to the underlying waveforms being different or whether they are similar waveforms buried in noise.

4.2.2. Semi-Similar Events

In the real world only a small percentage of events are truly repeats where the source waveforms are nearly identical. For example in China it appears to be about 10% (Schaff & Richards, 2004). For cross correlation methods to be applied on a broader scale it is necessary to extend the usefulness of these measurements to less similar events. This means we would like to be able to correlate events that are further apart, with less similar mechanisms, and greater magnitude differences to incorporate as many events as possible. Lowering the threshold of similarity comes with a cost. We still want it to be within the errors of the application. For making location estimates we basically want the correlation-measured delay times to be more accurate than the phase picks. If they are not then we can reject those thresholds. Most location work on a local scale uses CC thresholds of 0.6 or greater depending on the filter bands used and window length.

Here we show for an example that the error tolerance for detection work may be much less. The cyan colors between the two major high similarity clusters in Figure 11 show CC values of around 0.35. This is not a signal to noise problem since they correlate at 0.9 values with other events within the clusters. The cross correlation traces for one of these inter-cluster pairs are shown in panels 1, 3, 5, and 7 on Figure 18 and have values of 0.39, 0.39, and 0.35 for the vertical, north, and east components. The scaled cross correlation traces are shown in panels 2, 4, 6, and 8 and have values of 6, 7.2, 6.4, and 10.2. These are all above our trigger level for SCC of 6. Clear detection spikes can be seen on all the traces coming in at 50 s. Figure 19 shows the waveforms for the two events along with CC as a function of time in the seismogram for shorter windows. What can be seen is that portions of the seismogram are similar where the CC values exceed 0.5 and reach 0.75 for the shorter windows. Looking at those places on the actual waveforms verifies that they are more similar. However other portions of the Lg-wave are not similar.

One possibility for how these events may have semi-similar waveforms can be illustrated with a cartoon (Figure 20). Suppose two events occur in nearly the same spot recorded at a station denoted with the triangle. If the events have slightly different mechanisms, arrivals later in the coda may come out of different quadrants. In this case arrival "2" has a reverse polarity and would be dissimilar (or in this perfect example exactly anti-correlated). For ray paths "1" and "3" the waveforms are similar. Unfortunately we don't have focal mechanism data for these events in China. Hopefully these questions can be addressed with our work in northern California where we have high precision locations and abundant mechanism data. A better understanding is needed to see how variable the events can be and still provide useful detections.

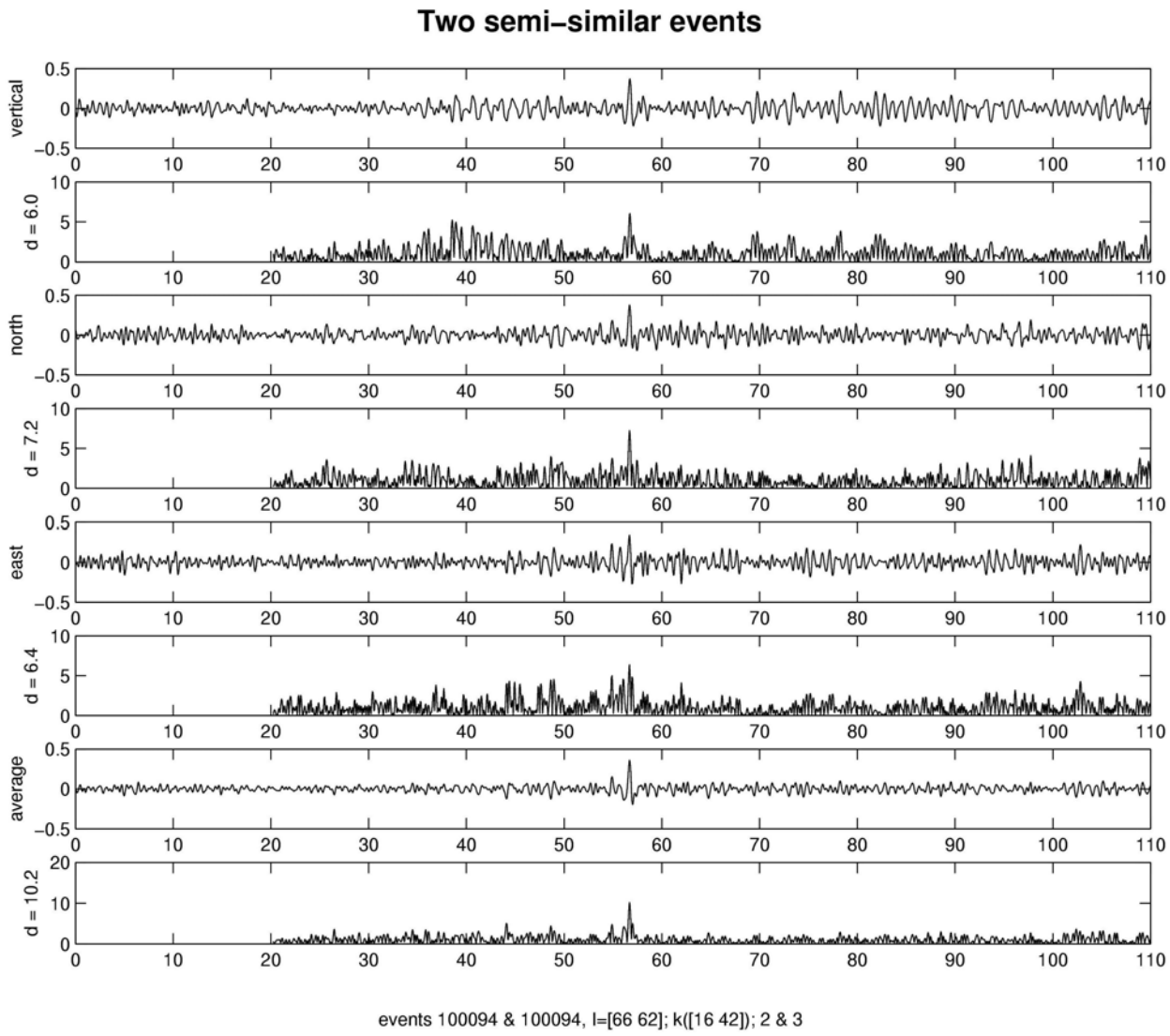

Figure 18. Cross correlation and scaled cross correlation traces for two semi-similar events.

4.2.3. Large and Small Event Correlations

Since we are presumably going to take a larger event as a master template to try to detect a smaller event in the noise it is helpful to know what range of magnitude differences will still correlate to provide satisfactory detections. Where the corner frequency is higher than the filter band used theoretically this is not as much of an issue. When the corner frequency is within the pass band, source finiteness is a factor that will change the shape of the waveforms and therefore the degree to which they correlate. Prior work has shown that events of similar magnitude tend to correlat well with each other (small with other nearby small events, and large with other large events even if the source

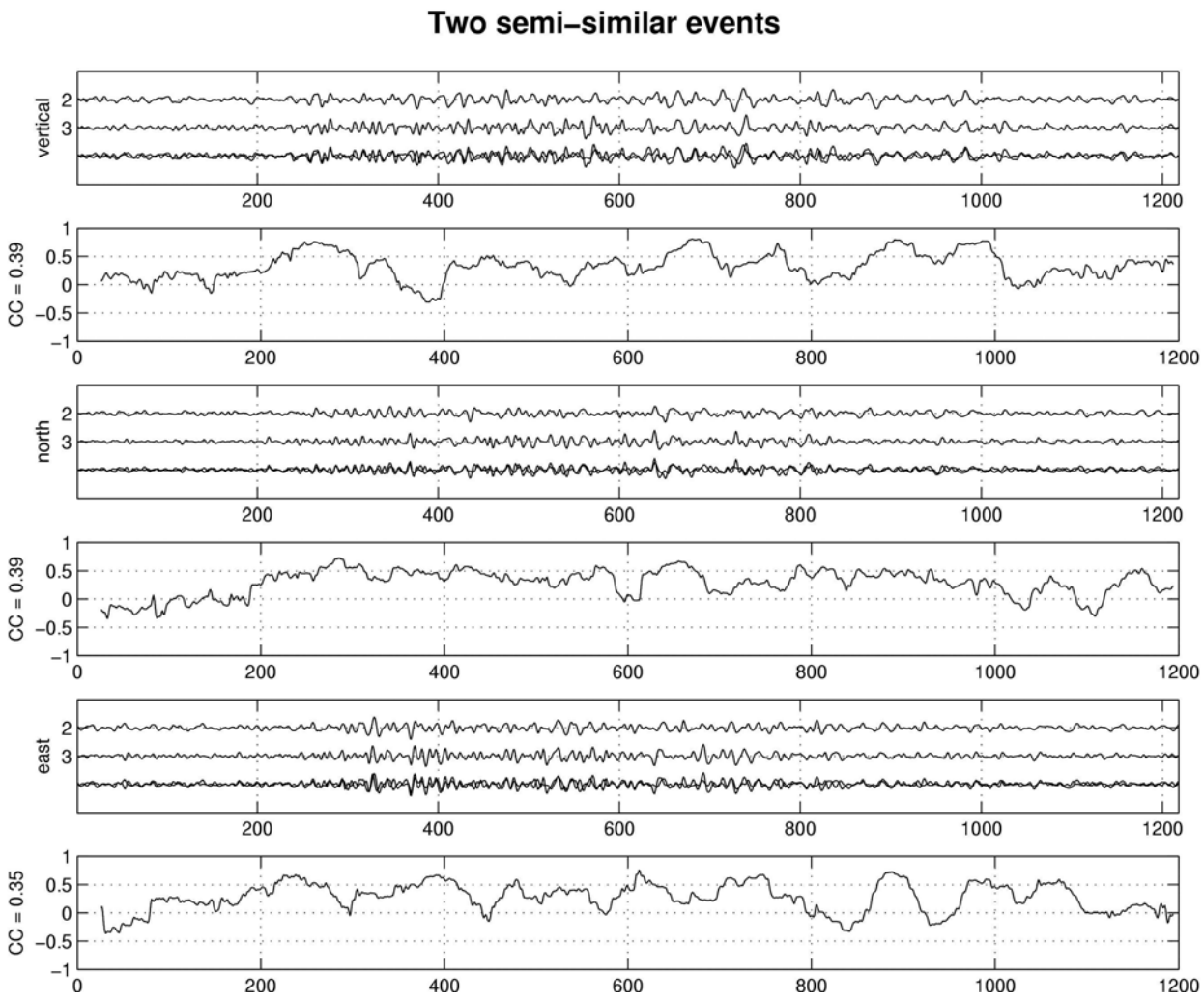

Figure 19. Waveforms for two semi-similar events on vertical, north, and east components. Bottom trace shows events superposed. Even panels show CC for a moving window of shorter length through the seismogram.

areas are not overlapping (Schaff et al., 2004)). Figure 21 shows an example of a magnitude 5.5 event that correlates with a 3.2 event at the 0.5 level (a 2.3 unit difference). The travel time is reduced to the Lg-wave arrival. Note that the Lg-wave is barely discernible to the eye for the magnitude 3.2 event. The first arriving P-wave (at -80 s) is well below the noise. Figure 22 shows the waveforms zoomed into the Lg-wave window that was correlated and with normalized amplitudes. The averaged 3-component cross correlation trace is a well-defined spike (not shown) because the value is quite high 0.5 for these window lengths and frequency content. On the superposed traces it can be seen that indeed several of the wiggles, peaks, and troughs do line up. Figure 23 shows another example from the same sequence of events at Xiuyuan where a magnitude 5.8 correlates with a magnitude 2.5 at the 0.25 level (an astonishing 3.3 unit difference or a factor of about 2000 in amplitude). At these levels it is harder to see the similarities in the

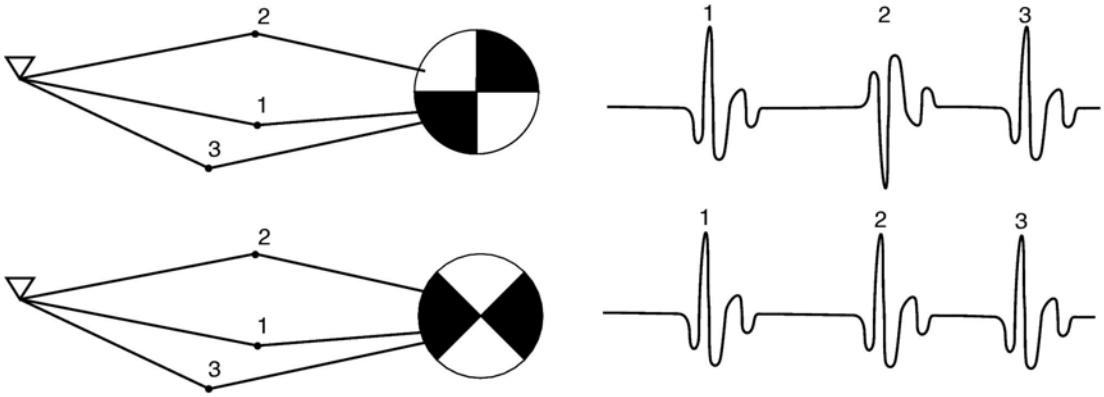

Figure 20. Possible explanation for how semi-similar waveforms are produce at a station. Two events with the same location but slightly different mechanisms have ray paths leaving different portions of the focal sphere.

waveforms by eye but the cross correlation traces show a clear detection spike on all three components that align to the nearest sample which is independent confirming evidence that the detection is real. And in this case we know it is real because the events are known with their corresponding Lg-wave windows. There is a need, however, to comprehensively study on a large scale the effect of source finiteness for less than perfect matches that still may be useable for correlation detection.

4.2.4. Buried Aftershocks

A second application of a correlation detector pertains to the issue of discrimination illustrated when a small seismic event occurred on August 16, 1997 near the former Soviet nuclear test site on the island of Novaya Zemlya. Because of its proximity to the test site serious concerns were raised as to whether this event was a clandestine nuclear test in violation of the Comprehensive Nuclear Test Ban Treaty (CTBT) signed 11 months previously by Russia. By taking a data window around the mainshock and cross correlating forward in time, convincing evidence was found of an aftershock occurring four hours later (Richards & Kim, 1997; Gibbons & Ringdal, 2006). Some large nuclear tests have had aftershocks, but not ones with this small a magnitude. Also a second seismic event can follow close in space and time to a nuclear test due to cavity collapse but the waveform for such an event is very different. The fact that this small seismic event had an aftershock with a similar waveform, taken alone, was very strong evidence that it was an earthquake.

We found two examples of a small aftershock detected seconds after a mainshock in our 1999 Xiuyan case study. They can be seen as the second blip on the vertical, north, east, and averaged CC traces in Figure 12 (1st trace) and 13 (sixth trace). These are new events discovered that aren't in the ABCE because they are in the coda of the mainshocks. Figure 24 shows the first one zoomed in. The waveforms are shown in the top three panels. The bottom panel shows the CC traces in red, blue, and green for the three components. The main shock detection spike clearly comes in around sample 1500.

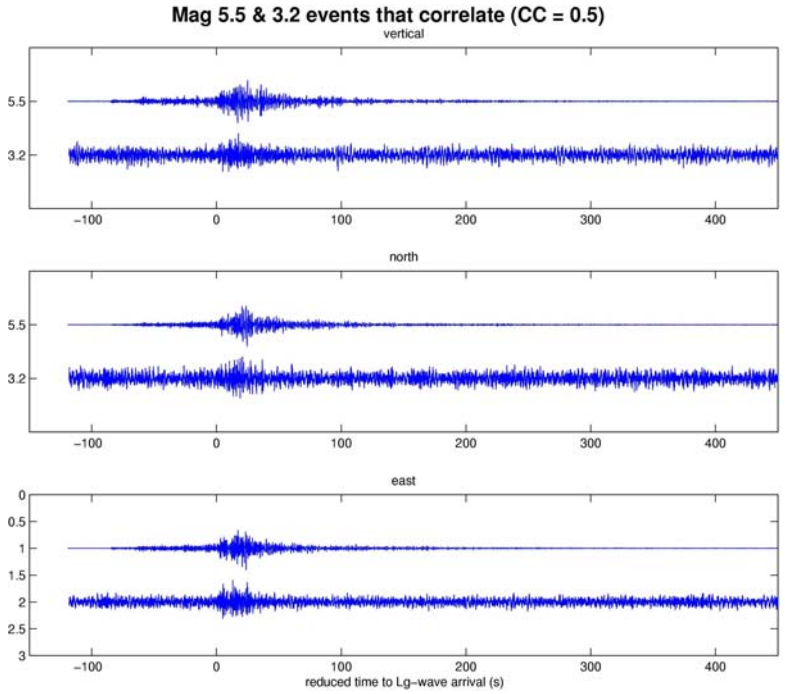

Figure 21. Waveforms for a magnitude 5.5 event and 3.2 that correlate with CC = 0.5. (Lg-wave at 0 s.)

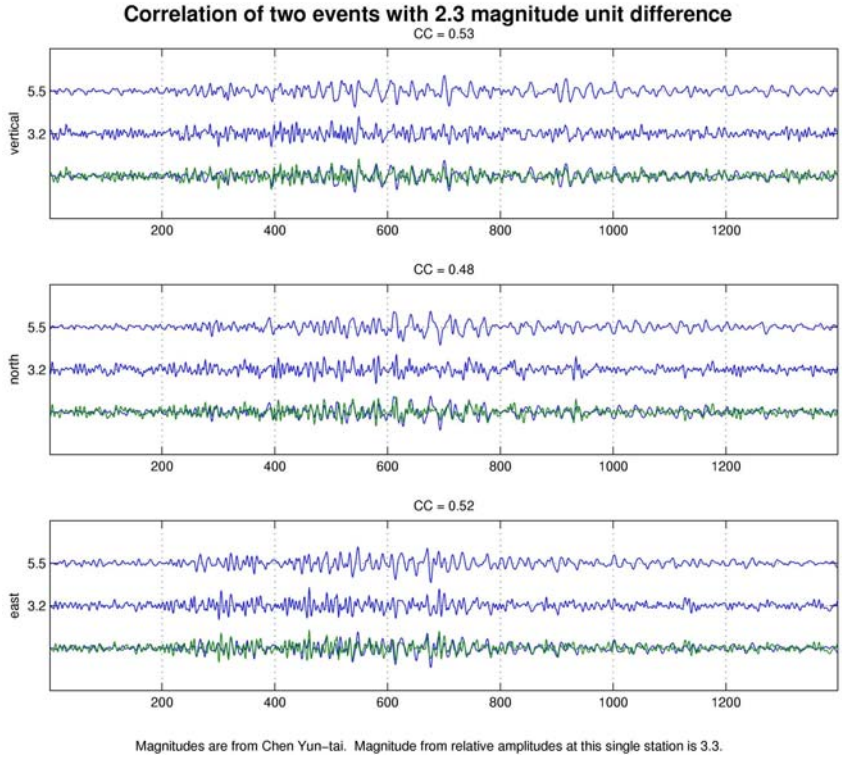

Figure 22. Zoom in on Lg-waveforms for Figure 21. Bottom trace shows the events superposed.

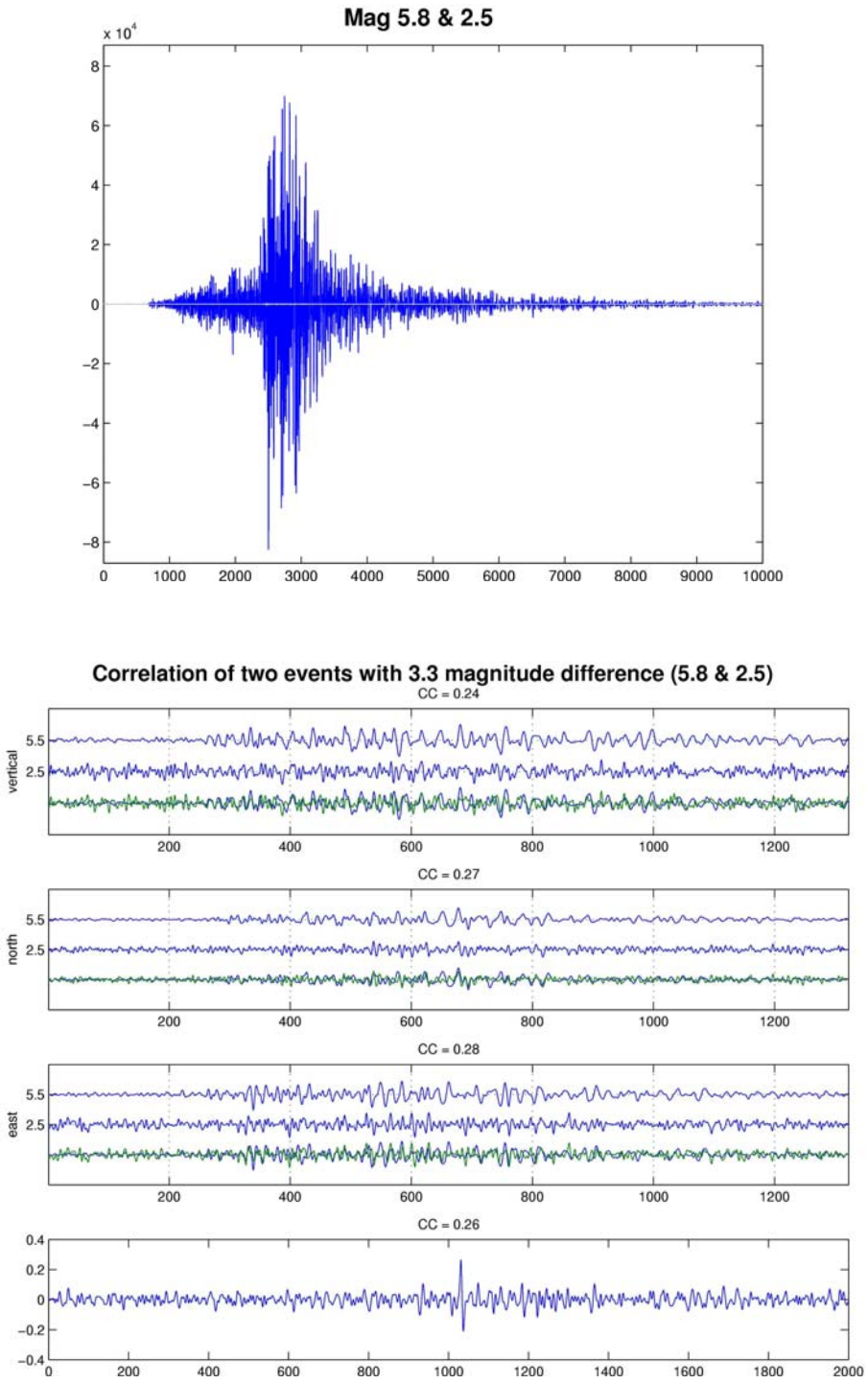

Figure 23. top) waveforms to scale for a magnitude 5.8 (blue) and 2.5 (gray) that correlate with each other. bottom) Normalized waveforms to show similarity. Bottom traces in each panel show pair superposed. Last panel shows averaged correlation trace with a significant detection spike.

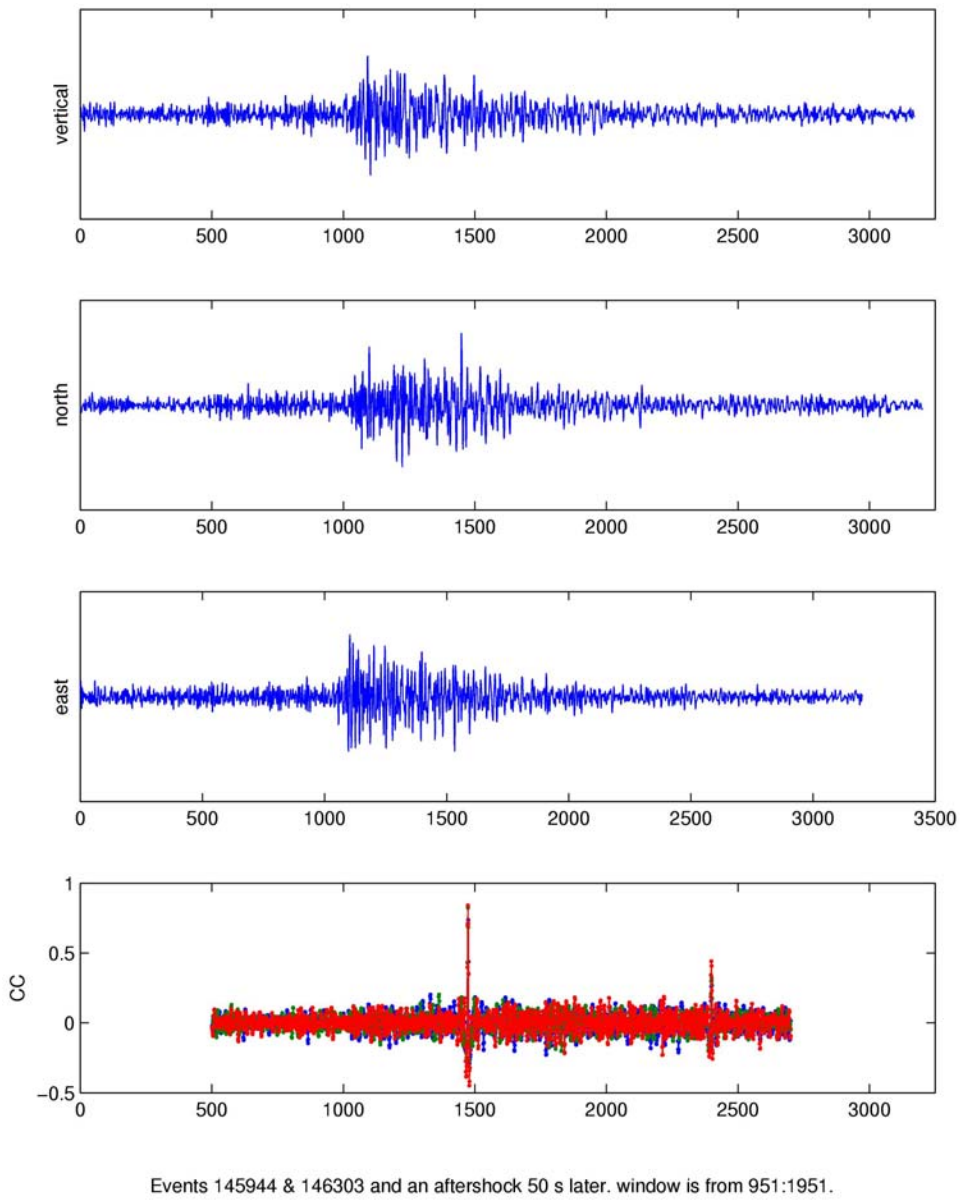

Figure 24. Example of an aftershock (spike at 2400 samples) detected after a mainshock (spike at 1500 samples) on three component data. Aftershock is not visible to the naked eye in the seismograms.

The aftershock detection spike around 2400 is also clearly seen on all three components aligning to the nearest sample (three independent tests confirming a true detection).

Figure 25 shows the waveforms for the master template aligned with the aftershock. CC values are 0.3, 0.32, and 0.44 for the three components. Several of the peaks and troughs are seen to line up by eye on the superimposed traces. The amplitude of the aftershock is about one third that of the main shock, corresponding to an event about 0.44 magnitude units lower.

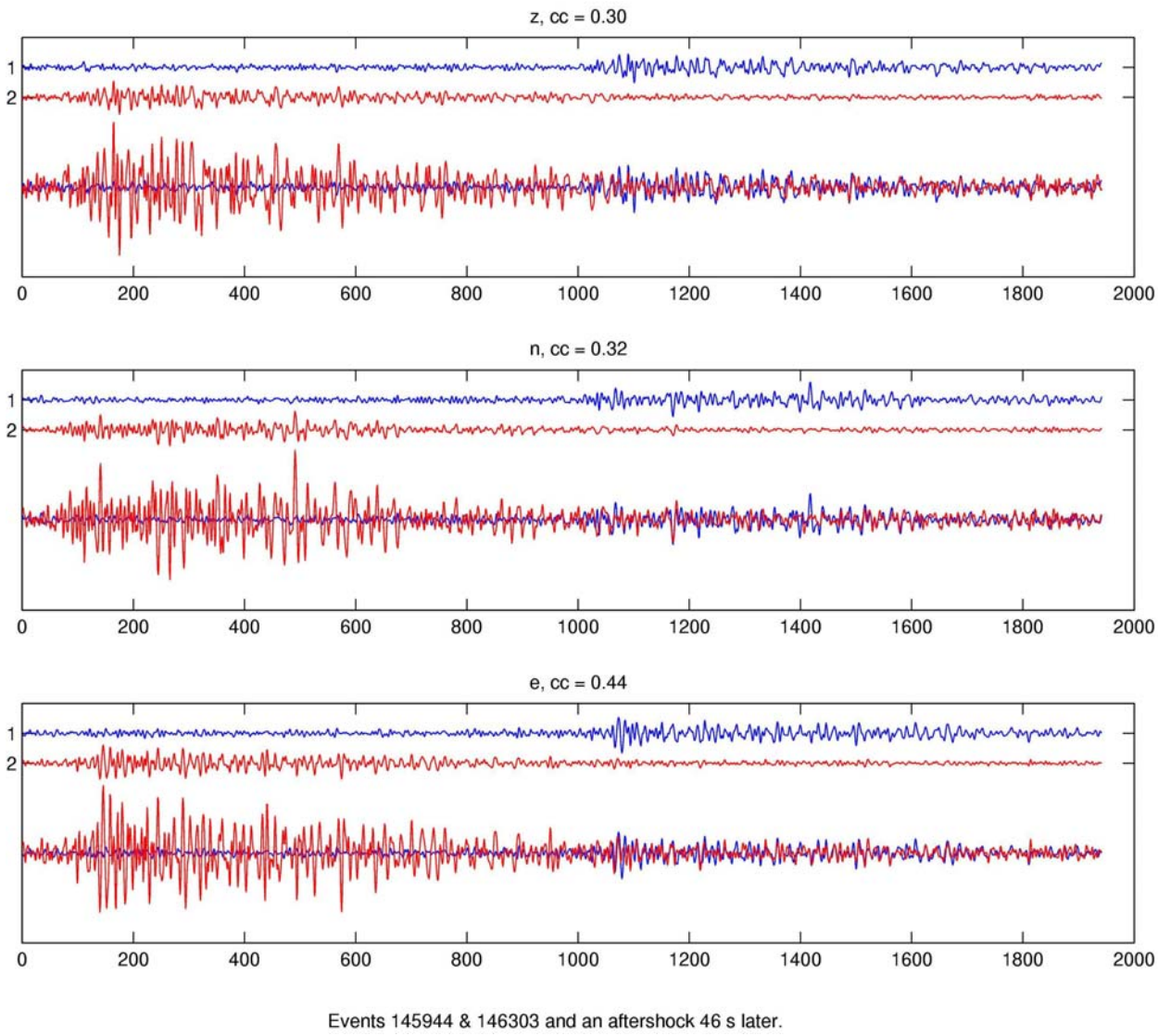

Figure 25. Alignment of master template (blue) with aftershock buried in the coda of a main shock (red) from Figure 12 for the vertical, north, and east components. Correlation coefficients for the aftershock are given in titles for each subplot. Third row in each panel scales the master event amplitude to the event in the coda for comparison of traces.

Detection of seismic events in the presence of other seismic signals is a problem because the background "noise" is higher. This was a difficulty during the days following the great M 9.0 Sumatra - Andaman Islands earthquake on December 26, 2004. There were so many intersecting signals traversing the globe it was hard to make detections and associations. Data centers were swamped like the NEIC and the IDC in Vienna and in the latter case even ceased bulletin production for several days. Correlation detectors may be able to extract more of the aftershocks in the days immediately following.

4.3. Large-Scale Application to China

The left plot of Figure 26 shows the 18,886 events in and near China and 363 stations used for this study. The events come from the Annual Bulletin of Chinese Earthquakes (ABCE) from 1985 to 2005. The stations are those for which waveforms are available at the IRIS DMC. Several of the stations are only temporary deployments. There are only a few long-running stations in China, which correlation techniques work best for, so the actual network of stations for most of the events is quite sparse with large inter-station distances. A total of 111 million cross correlations were performed taking about 2 weeks of continuous processing time on a four CPU computer. All events with separation distances less than 150 km were correlated. In other words, every event is a master event to see the maximum possible number of correlations possible. So far only Lg-phases have been processed. 50 s windows were used searching forward and backwards 30 s using time-domain cross correlation. The seismograms were filtered from 0.5 to 5 Hz. The cross correlation traces for the three components are averaged together to constructively enhance the detection spikes when present. A "scaled cross correlation coefficient" (SCC) was used to initially sift the data. All correlations with SCC > 4.5 were saved.

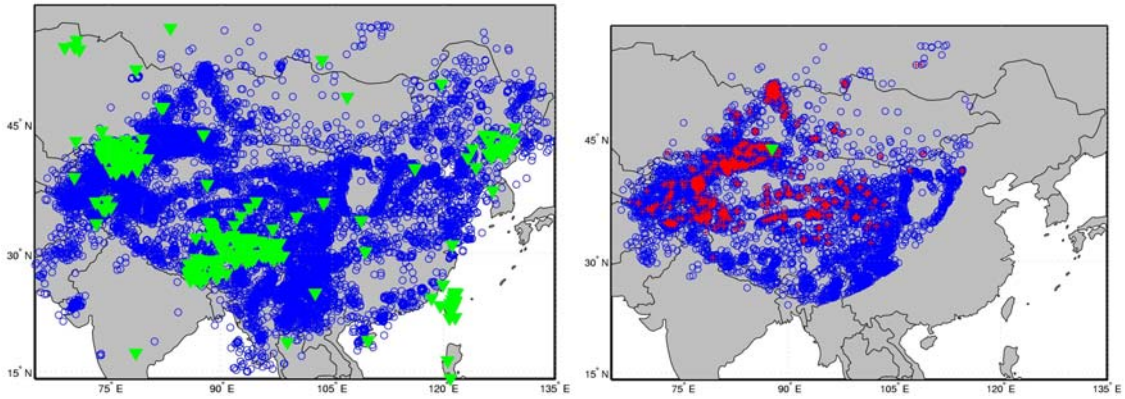

Figure 26. (left) 18,886 events (blue circles) recorded at 363 stations (green triangles) in and near China. (right) events in blue recorded at station WMQ (green triangle). 17% of the events (red) have CC > 0.5 with at least one other event at this station.

A study was made to estimate false alarms using real data for selected stations. It is assumed that randomly selected time windows should not contain events that correlate. From this it is possible to determine curves of the false alarms per day as a function of cross correlation coefficient (CC) and SCC. For traces that are already sifted with an SCC 6, a CC of 0.24 corresponds to approximately one false alarm per day. A CC of 0.5 corresponds to 0.0022 false alarms per day. An SCC of 6.65 corresponds to a little over one false alarm per day. Based on the pair-wise distance matrix for the correlations and

the number of samples in the lags searched over we can estimate the percent of the 18,886 events for which detections are processed that would be expected to be false alarms. Because a CC of 0.5 appears to be so robust for these long 50 s windows we assume that a trigger at a single station provides a reliable detection. The right plot of Figure 26 shows the events within 20 degrees recorded at station WMQ. The events in red are 17% of the total that have CC > 0.5. Notice how the Tibetan Plateau blocks the Lg-wave propagation for this station. For the other thresholds of one false alarm per day, it is necessary to require at least two stations observing that pair to count as a detection. Using a combination of these criteria (CC > 0.5 at 1+ stations, CC > 0.24 at 2+ stations, or SCC > 6.65 at 2+ stations) we estimate the percent of events that would be false alarms to be approximately 3%. Applying these criteria to the time windows corresponding to where the theoretical travel times would arrive for the 18,886 events results in 12,902 events that are detected or 68%. Therefore it is estimated that 65% represent true detections.

Figure 27 shows the magnitude distribution for the 18,886 events and the 12,902 events detected by cross correlation. 8,358 events found by a "pIDC" type detector are also shown (44%). The pIDC employed an STA/LTA detector on the vertical component for P-waves in overlapping narrow pass bands. Three stations triggered are necessary to be counted as a detection. The number in parentheses in the legend is the 95% confidence lower limit of the magnitudes detected. Overall it can be seen there is a 0.2

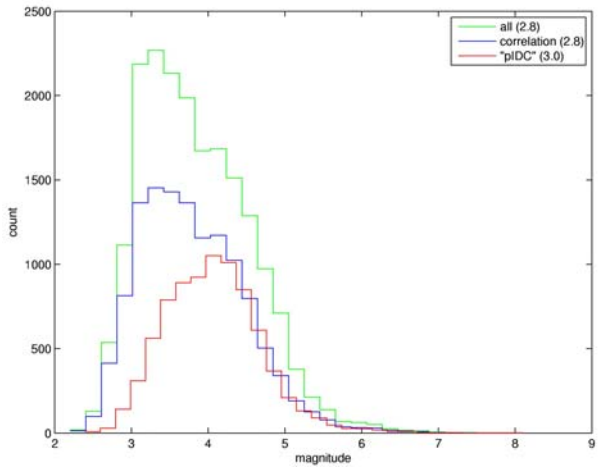

Figure 27. Histograms of the magnitude distribution for all 18,886 events and the 12,902 events detected by correlation and 8,358 found with pIDC type procedures. Correlation finds more events and lower magnitude thresholds restricted by the lower limit of the catalog magnitude of completeness. The number in parenthesis in the legend gives the 95% confidence lower limit for the magnitude distribution.

unit reduction in magnitude detection threshold for correlation compared to the "pIDC". However 2.8 is also the 95% lower limit of all the magnitudes in the catalog. Therefore the range of magnitudes is not low enough to test if correlation improves detection thresholds more than 0.2 units overall.

More insight can be gained by plotting magnitude as a function of station distance. The left plot of Figure 28 shows the detections for the "pIDC". The red line is the 95% lower limit of the magnitudes in 50 km bins of station distance. The trend is seen to increase from about 2.9 at zero to 3.5 for 2200 km station distances. The right plot of Figure 28 shows that the number of observations that are detected steadily decreases after about 500 km. Figure 29 shows similar plots for the correlation detector. The red line is the 95% lower limit of the magnitudes in 50 km bins. The green line is the 95% trend

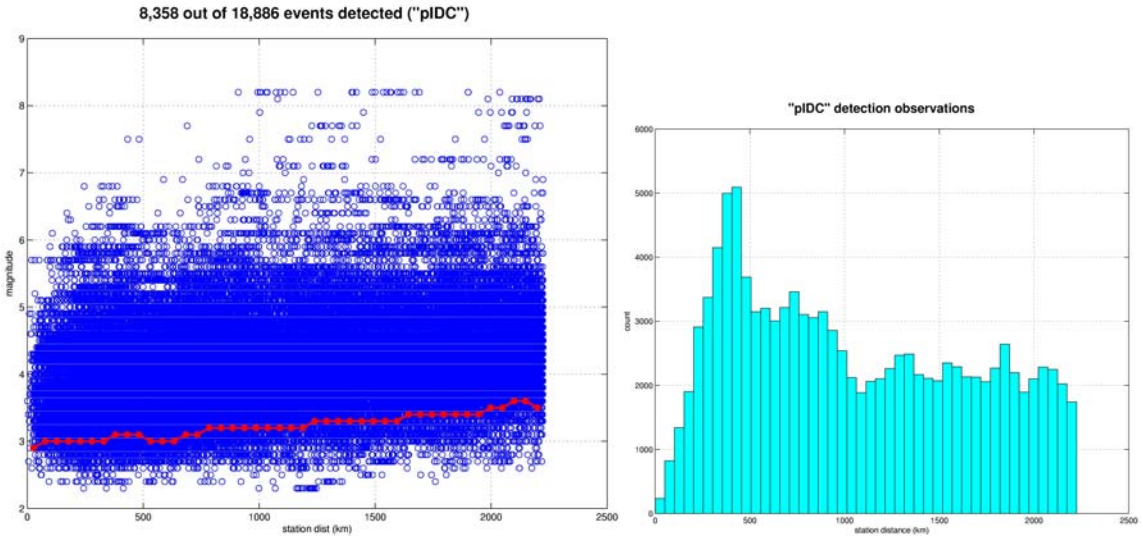

Figure 28. (left) Plot of magnitude vs. station distance for all triggers for "pIDC". Red line shows 95% confidence lower limit of magnitude in 50 km bins. (right) Histogram of number of observations as a function of station distance.

from Figure 28. It is seen the two lines diverge for greater station distances. The right plot of Figure 29 shows that correlation detection observations drop off faster with station distance. The difference between the two 95% trends is plotted in Figure 30. At zero the difference is 0.2 units and then increases to a maximum of 0.9 units at greater station distances. The interpretation of these results is that for longer station distances the magnitudes in the catalog are sufficiently complete to observe nearly an order of magnitude improvement in threshold reduction between the two techniques. For closer station distances lower magnitude events are not available in this catalog to test if the full unit reduction still holds.

We see that approximately two thirds of the events in the catalog are detected by cross correlation (12,902), a sizeable fraction indicating that these methods can be applied on a broad scale across diverse tectonic regions. However, this catalog was based on a much denser network of stations than the one that we have waveforms available for. To get an idea of how applying correlation methods to an existing network would improve things it is instructive to see how many events are detected by correlation and an

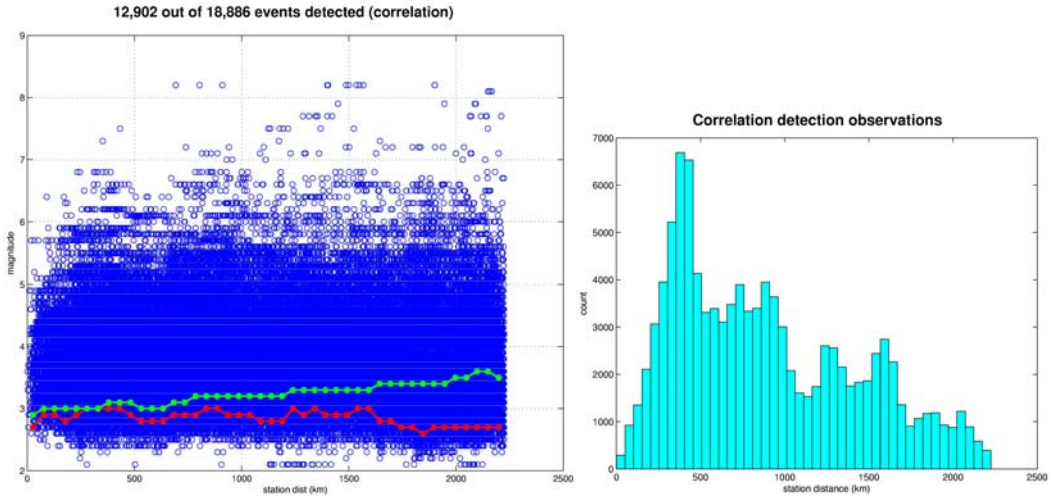

Figure 29. (left) Plot of magnitude vs. station distance for all triggers for correlation detector. Red line shows 95% confidence lower limit of magnitude in 50 km bins. Green line is the curve for the "pIDC" from Figure 3. (right) Histogram of number of observations as a function of station distance.

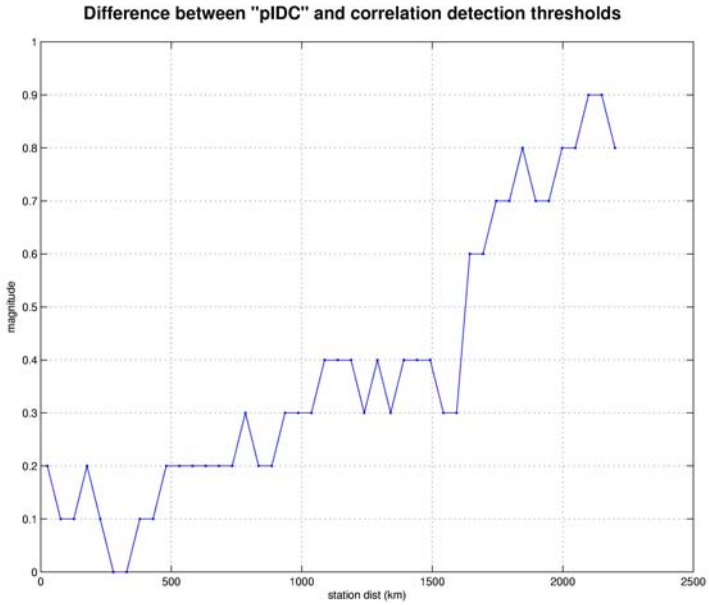

Figure 30. Difference between "pIDC" (green) and correlation (red) lines on Figure 4 as a function of station distance

STA/LTA detector on the same network. 7,063 events were detected by correlation out of the 8,358 found by a "pIDC" procedure or 85%. For comparison Schaff and Waldhauser (2005) determined that 95% of the 225,000 events in northern California correlated at four or more stations with CC > 0.7 with at least one other event. Therefore

correlation is able to detect the great majority of the seismicity for these large regions of seismicity. This can be an important independent confirmation of the existence of new events and help to weed out false alarms. Besides lowering magnitude detection thresholds correlation also detects more events that the "pIDC" procedure missed due to a variety of reasons (Figure 1 right plot). The correlation detector finds 5839 additional events over the 8,358 events from the "pIDC" detector or a 70% increase. Therefore we might expect catalogs for existing networks to also increase with the complementary benefits of correlation detector techniques.

4.4. Large-Scale Application to Parkfield, California

5076 events at Parkfield, California, in red on Figure 31 were processed in a similar manner as for China at seven continuously operating stations (blue triangles). This time, however, P-, S-, and Lg-waves were all analyzed with 15 s, 20 s, and 50 s windows respectively and 10 s, 15 s, and 30 s lags searched over respectively. The filter bands were 0.75 to 2 Hz, 0.5 to 3 Hz, and 0.5 to 5 Hz respectively. All pairs within 6 km were considered amounting to 53 million correlations and about 2 weeks of continuous processing time on a four processor machine.

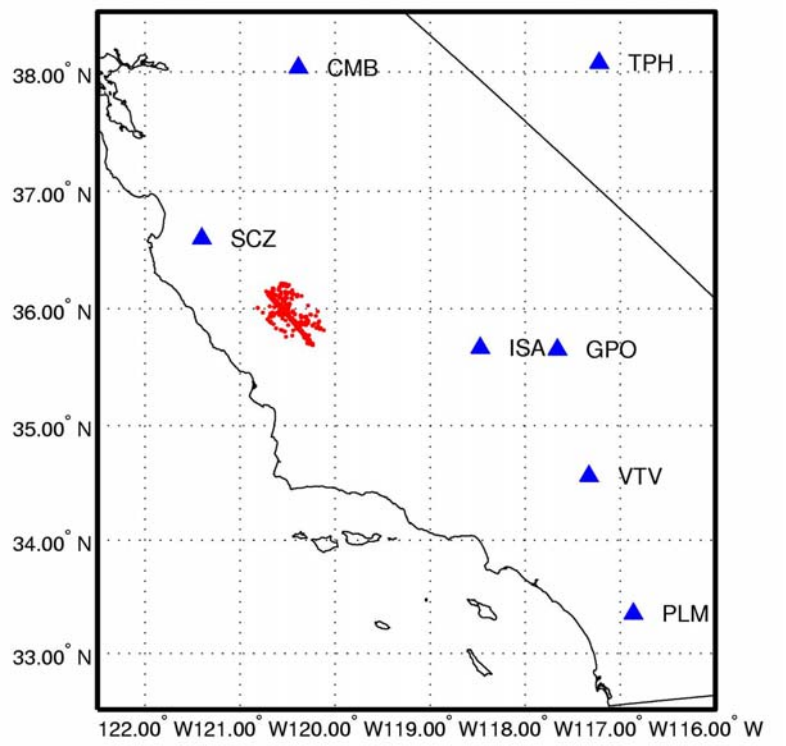

Figure 31. 5076 events at Parkfield (in red) processed at seven stations (blue triangles) for a correlation and standard detector.

We empirically estimate false alarm rates by running the codes in an identical manner with the same parameters except for shifting the windows 120 s before the expected P-wave arrival. The idea behind this is that the windows should contain only noise and so any trigger for a given threshold would then be considered as a false alarm. This is the most robust method we know of for estimating false alarms for our comparisons since all the stations and event pairs are the same and the windows are centered on the noise characteristic right before the signal comes in. Table 2 shows how we determine the number of true detections for each station and phase using this procedure. For example, the P-waves at station SCZ had 4096 observations with SCC >= 6 for time windows centered on the expected signal arrival. Comparing this to 623 observations with an SCC >= 6 for the same processing except on noise allows us to estimate the number of true detections as the difference between the two or 3473 (Table 2).

Table 2. Number of true detections with SCC >= 6

Station	Phase	Detections
SCZ	P	3473
SCZ	S	7658
SCZ	Lg	10612
ISA	P	1225
ISA	S	6228
ISA	Lg	12986
CMB	P	720
CMB	S	3002
CMB	Lg	3787

Station	Phase	Detections
VTV	P	-37
VTV	S	23
VTV	Lg	404
TPH	P	27
TPH	S	268
TPH	Lg	515
PLM	P	13
PLM	S	281
PLM	Lg	479

From this table of event pair detections we next plot the distribution in Figure 32 of the inter-event separations for the event pairs that are detected with SCC >= 6 and compare it to the input inter-event separations searched over (in this case it was an earlier run considering separations out to 20 km instead of 6 km like indicated above). From the left panel of Figure 32 it is clearly seen that most detections occur within a kilometer of each other. This is independent confirmation that these are mostly true detections because the false alarms would tend to reproduce the distribution of the input observation event pair matrix on the right panel of Figure 32 which is not what is observed.

The stations in Table 2 are listed in order of increasing station distance. It is observed that the two closest stations give the most detections. It is also seen that the Lg phase gives the most detections even though it doesn't propagate as efficiently in California. Further away stations produce relatively few true detections compared to the number of false alarms. Therefore we decide based on this initial screening of all seven stations and

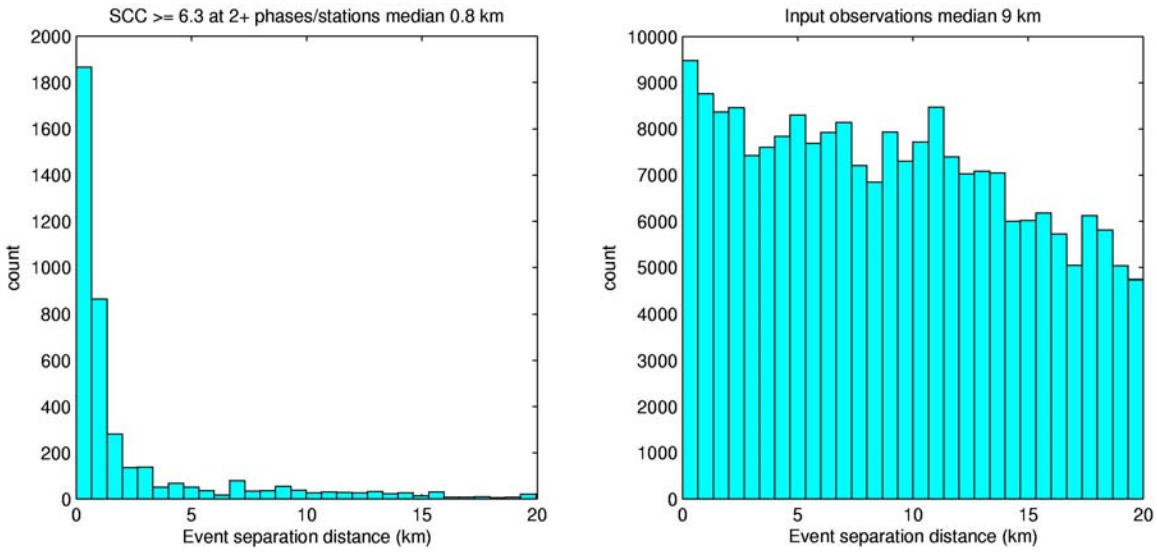

Figure 32. Inter-event separation for detected pairs on left and input matrix on right.

phases to use only Lg-waves at stations SCZ and ISA. We further require a selection criteria of event separations less than 1 km, lags searched over to be 0.3 s, and SCC >=5. We do this to maximize the amount of detections while trying to minimize the influence of false alarms. If we use selection criteria similar to what we used for China, however, with SCC >= 6, all event separations, and 30 s lags the overall statistics are not grossly different. Because we have good control on the locations at Parkfield with our high resolution catalog we'd like to see how well the correlation detector can perform. Lags of 0.3 s correspond to as much as 1 km relative location error for a group velocity of 3.5 km/s for the Lg-waves which is still a rather conservative estimate. Using 30 s (two orders of magnitude larger) is then overly conservative.

Using these selection criteria we find that 1357 events out of the 5076 or 27% are detected. The false alarms using these same criteria are 1.4%, so we estimate that the true detections are approximately 25.6%. We compare with the same "pIDC" procedures as used for China except counting a trigger if it occurs within 5 sec of the first arriving P-wave this time since the locations are more accurate. The "pIDC" procedures find 140 events out of the 5076. Figure 33 shows the magnitude distributions. It can readily be seen that the correlation detector finds approximately 10 times the number of events that the "pIDC" detects. The correlation detector also finds 120 out of the 140 events that the "pIDC" detector finds or 86%.

The 95% confidence lower limits for the distributions in Figure 33 are: all (0.8), correlation (1.1), and "pIDC" (1.4). This represents a 0.3 magnitude unit reduction between the two techniques which is not as great as expected from prior work where a full unit was measured. This time, however, the completeness of the catalog at Parkfield is not the reason like we supposed it was for China. Further examination brings up some issues and explanations. If we assume that a 1.4 lower limit is really representative of a 95% confidence level we see that this gives 132 events for the "pIDC" out of a possible 1344 in the catalog or 9.8% which is far below 95% level and immediately presents a

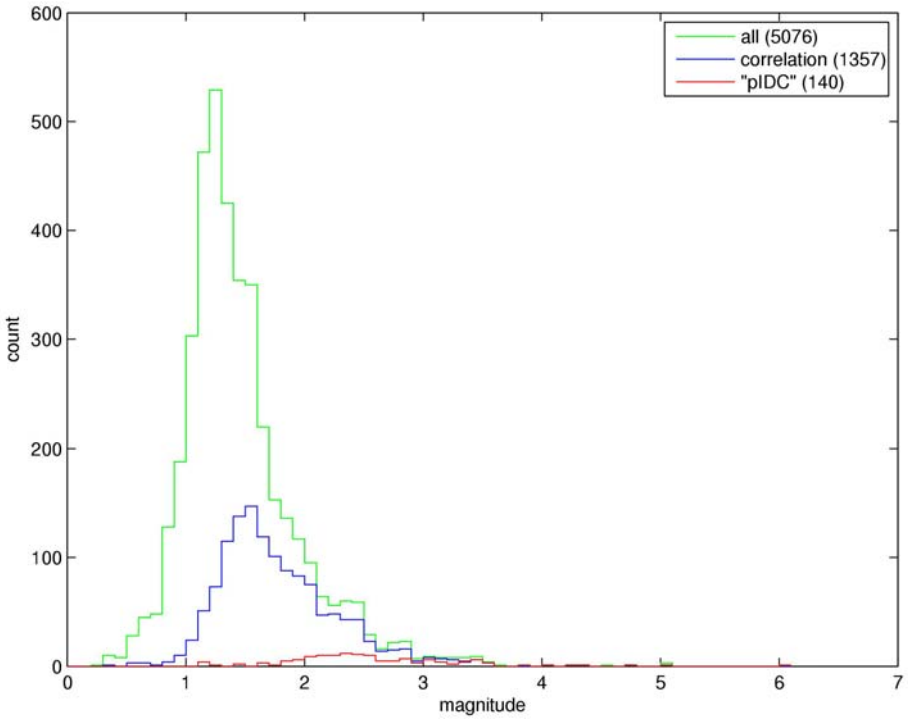

Figure 33. Magnitude distribution for Parkfield catalog and correlation and "pIDC" detectors. Number in parentheses shows total events.

problem. Similarly, if we zoom into the "pIDC" distribution more closely in Figure 34, we see that 22 events out of 2581 in the catalog have magnitudes less than 2 or 0.85%. A false alarm rate as low as 1% therefore could be the cause for these lower magnitude events that are "detected" in the "pIDC". The reason is that the Gutenberg-Richter magnitude-frequency relationship means there are orders of magnitude more smaller events than larger events. Therefore for a given false alarm rate it is more likely that the false alarm will be a lower magnitude event than a larger magnitude. This causes a problem for determining lower limit magnitude detection thresholds especially for the case when the total number of events is small (140 for the "pIDC) compared to the background catalog (5076). So if 22 events out of the 140 in Figure 34 are actually false alarms that is 16% of 140. So a 95% threshold would include false alarms. One idea for a solution is to use a different threshold (e.g. 100%-16% = 84%). If we consider a 90% lower limit than correlation has 1.2 whereas "pIDC" is 1.8 and therefore the reduction in threshold is 0.6 units. An 85% confidence lower limit has correlation (1.3) and "pIDC" (1.9) with a threshold reduction of 0.6 unit. This is closer to what is expected and presumably reflects the improvement more accurately since they are not as influenced by the presence of false alarms.

We explored a more representative measure of detection thresholds that takes into account the predominance of smaller magnitude events. Looking at Figure 33 again a more intuitive limit may be the first value where the detector finds 50% or more of the available events within a given magnitude range. This would avoid the problem seen

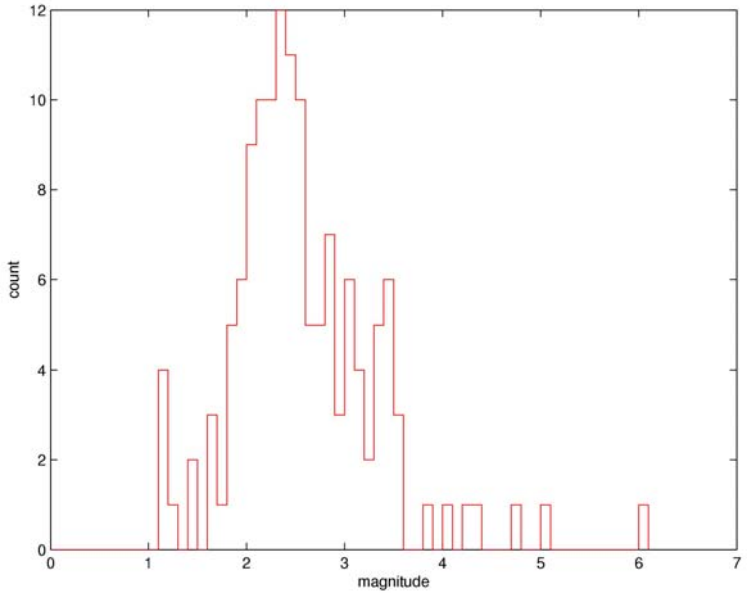

Figure 34. Zoom in of "pIDC" magnitude distribution from Figure 33.

before that a 95% threshold actually only captured 9.8% of the events. We can create a graph of what we term a normalized probability density function (PDF) in Figure 35 by taking the red "pIDC" detector curve of Figure 33 and dividing it by the green curve for the entire catalog. This changes the number of detections into percentages that aren't skewed by the large numbers of small events. If all the events were detected in the catalog then the PDF would be flat at unity (100% detection) across all magnitudes.

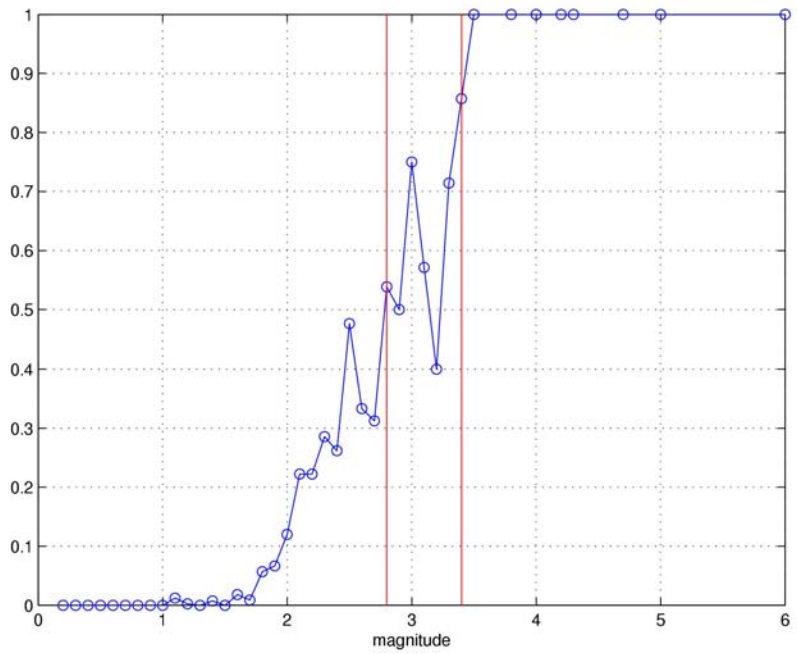

Figure 35. Normalized probability density function for the "pIDC" computed as the red curve from Figure 33 divided by the green curve.

The left red line on Figure 35 indicates the lower magnitude threshold of 2.8 corresponding to the first point at which 50% of the background catalog is detected in that magnitude bin. The right red line on Figure 35 marks where at least 80% of the events in the catalog are detected at 3.4. Figure 36 shows the normalized PDF for the correlation detector by taking the blue line from Figure 33 and dividing it by the green curve.

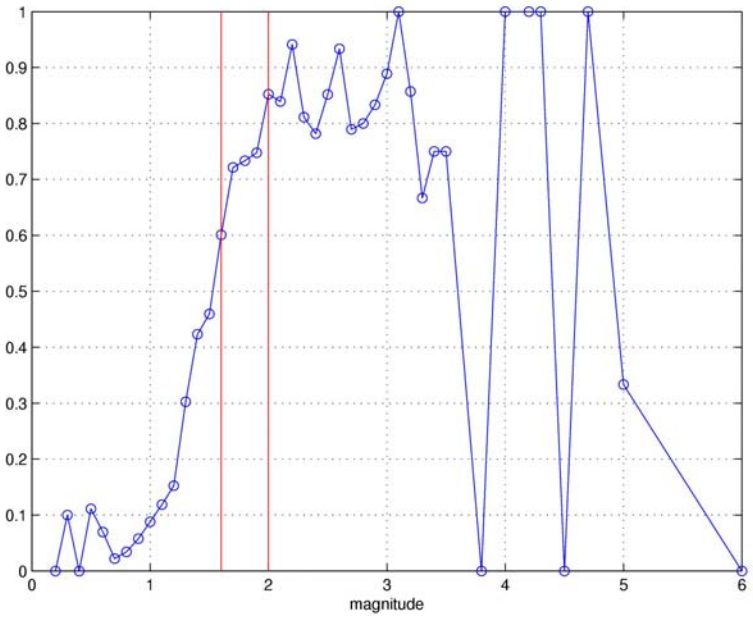

Figure 36. Normalized probability density function for the correlation detector computed as the blue curve from Figure 33 divided by the green curve.

The 50% red line on Figure 36 corresponds to a magnitude bin of 1.6 and the 80% red line falls on the 2.0 bin. The difference between the two detectors for this detection threshold measure is a reduction of 1.2 units for the 50% level and 1.4 units for the 80% level. This improvement is consistent with what we observed before with a full magnitude unit reduction in the previous studies. It also is a more representative and intuitive measure since a 50% value does correspond to the point where half of the earthquakes in the catalog at that magnitude range are actually detected instead of a 95% confidence limit giving the misleading information that 95% of the events are detected whereas only 9.8% were. Additionally by normalizing according to the PDF of the background catalog accounts for the Gutenberg-Richter magnitude-frequency relationship and reduces the impact of false alarms at lower magnitudes since a percentage is considered instead of a number.

An alternative way to measure the detection threshold is to convert the normalized PDF to a normalized cumulative density function (CDF) in Figures 37 and 38. The CDF is computed as the cumulative sum of the PDF and then normalized to one. A 95% confidence lower limit on the normalized CDF for the correlation detector corresponds to magnitude 1.3 on Figure 37. For the “pIDC” detector the 95% confidence lower limit corresponds to a magnitude 2.2 and so the reduction in threshold is 0.9 units using the normalized CDF as a measure. At the 90% confidence lower limit the values are 1.5 for

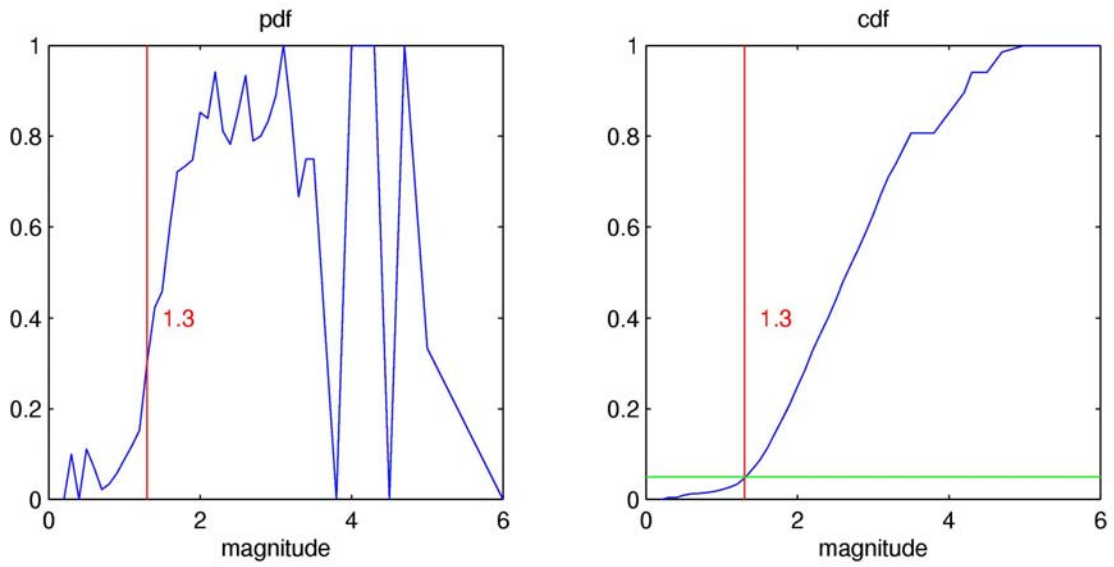

Figure 37. (left) Normalized PDF for the correlation detector from Figure 36. (right) Normalized CDF computed from PDF as the cumulative sum normalized to one. Magnitude 1.3 corresponds to the 95% confidence lower limit for the CDF.

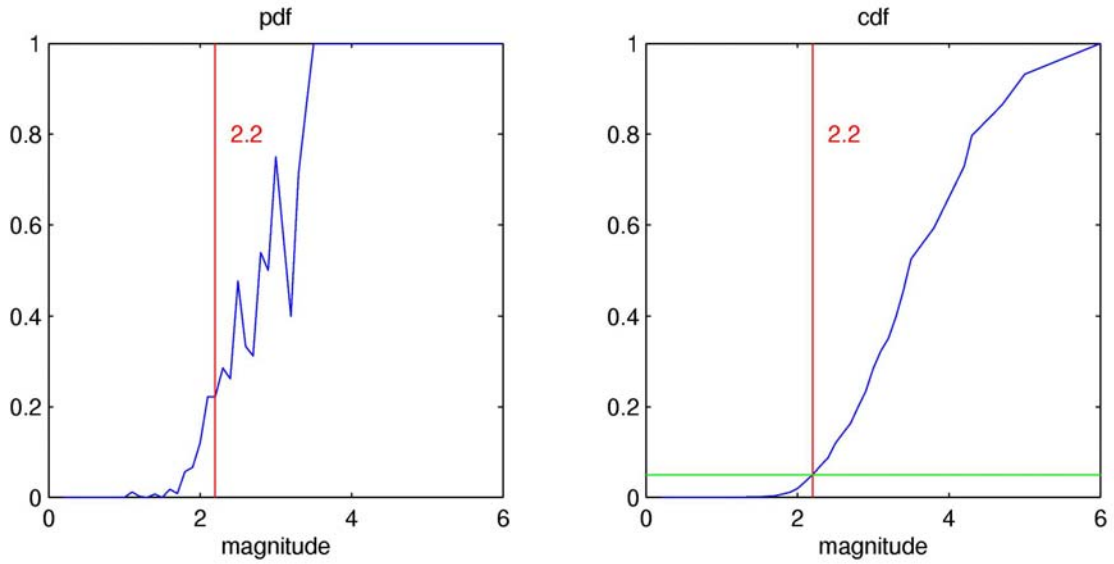

Figure 38. (left) Normalized PDF for the "pIDC" detector from Figure 35. (right) Normalized CDF computed from PDF as the cumulative sum normalized to one. Magnitude 2.2 corresponds to the 95% confidence lower limit for the CDF.

the correlation detector and 2.4 for the "pIDC" detector with a reduction in threshold again of 0.9 units. It seems that the normalized CDF is more robust of a measure of threshold since it was a 0.9 reduction for both 95% and 90% confidence limits whereas

the normalize PDF ranged from 1.2 to 1.4. The CDF curves are also smoother than the PDF curves. Finally a magnitude threshold that captures a normalized percentage of all earthquakes above a certain value instead of within a specific magnitude range seems more robust and less sensitive to the details of the magnitude distribution of the background catalog.

From Figure 37 it appears that M 1.3 is a good measure of the detection limit for the correlation detector for these stations at Parkfield. Figure 39 shows all the events in the high resolution catalog at Parkfield with M > 1.3. Events in red are detected by cross correlation and events in blue went undetected. It can be seen that the detected events are spatially well-distributed across the fault and at all depths. They tend to occur in tightly clustered areas. The undetected events occur in more diffuse areas. Figure 40 shows that the density of events within 1 km around detected events above M 1.3 is greater than that around the undetected events which may partially explain why they went undetected. The reason why the larger events were not detected is because of the inter-event distance

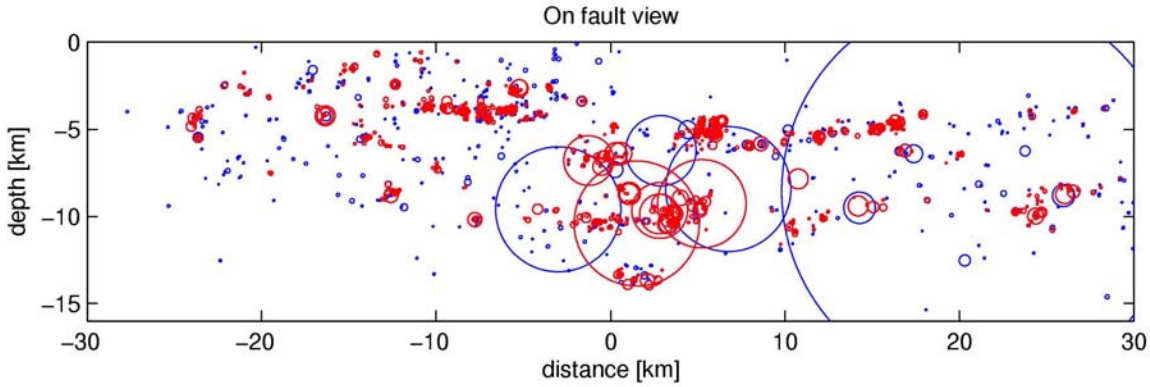

Figure 39. Events greater than M 1.3 in the Parkfield catalog. Red are detected by cross correlation and blue are undetected.

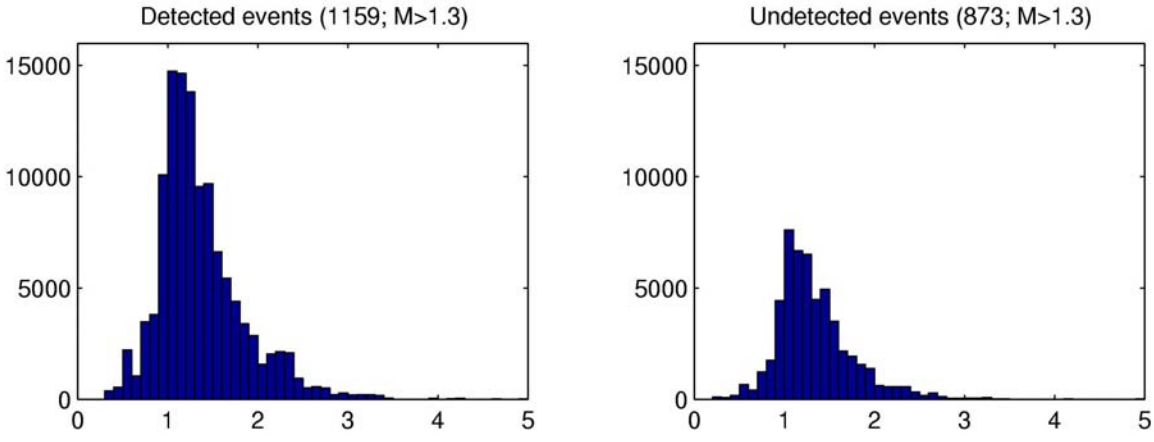

Figure 40. Density of events above M 1.3 as a function of magnitude for correlation detected events (left) and undetected (right).

thresholds being restricted to 1 km. A prior run with separation out to 20 km captured all of these larger magnitude events.

Figure 41 displays the distribution of magnitudes for the master events for which correlation detections were made. The master event of a pair is taken as the larger of the

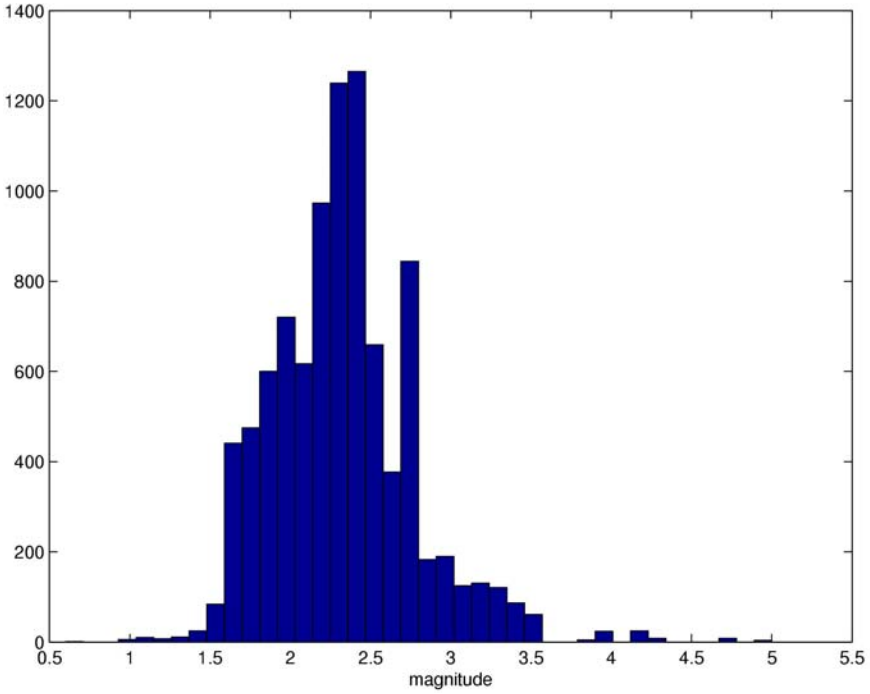

Figure 41. Distribution of magnitudes of master events for correlation detections.

two events. It can be seen that there is a sharp cut-off around M 1.5. From our semi-empirical synthetic runs we expect a full magnitude unit reduction in detection threshold for similar signals if the master event has no noise and the slave event contains noise. Therefore the lower magnitude limit for the correlation detector of 1.3 from Figure 37 for the slave event would correspond to a magnitude 2.3 master event or larger (one full unit higher which would have a good SNR). However for a magnitude 1.5 master event the SNR would not be as good. Therefore a M 1.5 master event would not be expected to be able to detect a M 1.3 slave event but rather another M 1.5 event.

Figure 42 plots the distribution of the magnitude differences for the detected event pairs and the input observation matrix. The two distributions are virtually indistinguishable and many pairs with separations larger than 1 magnitude unit are detected. This is a desirable feature since we hope to detect smaller events with larger events and also include detections for less than perfect waveform matches due to source complexities. For detection purposes it appears that larger magnitude separations are possible than for location work involving correlation measurements.

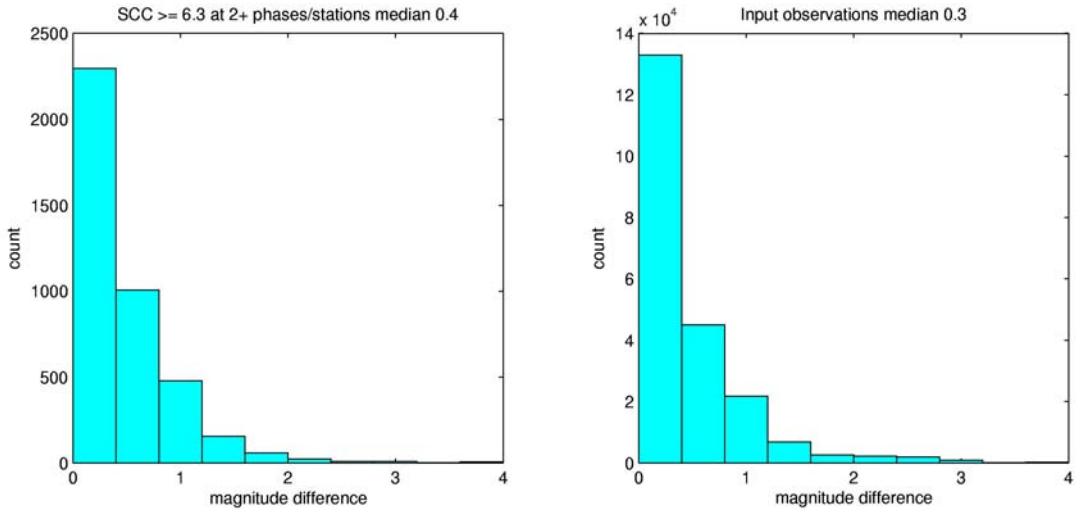

Figure 42. Distribution of magnitude differences for correlation detected event pairs (left) and the input observations (right).

For comparison we revisit the large-scale China results using the more representative measures of detection thresholds. Figure 43 displays the normalized PDFs for the correlation and "pIDC" detectors for the curves in Figure 27 for China. The 50% detection levels are 2.1 for correlation and 3.7 for the "pIDC" amounting to a 1.6 unit reduction in threshold. For the 70% level the values are 2.4 for correlation and 4.2 for "pIDC" corresponding to a 1.8 reduction in threshold.

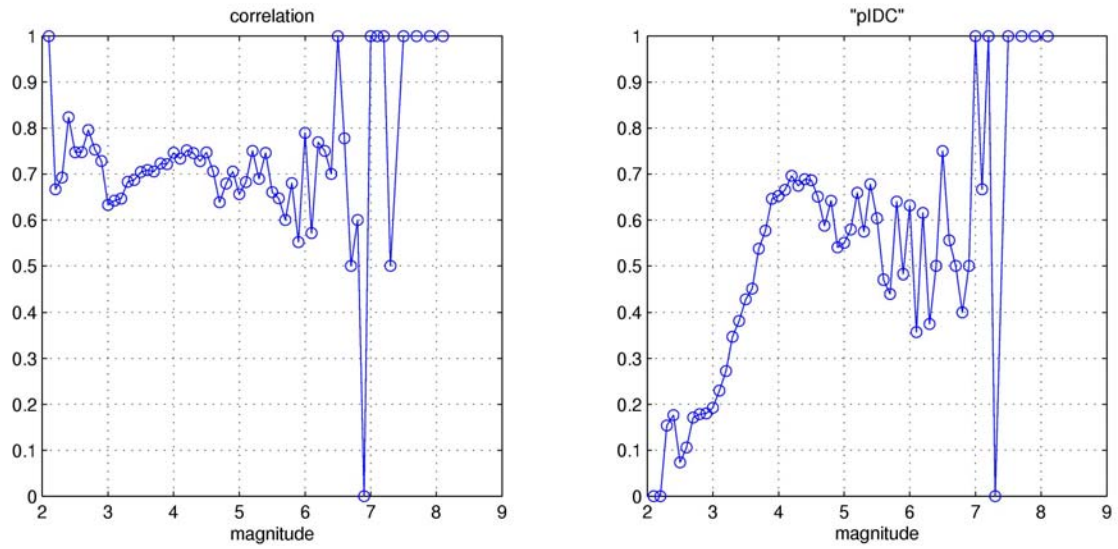

Figure 43. Normalized PDFs for correlation detector for China (left) and "pIDC" detector (right).

Similarly the normalized CDFs for China are shown in Figures 44 and 45. The 95% confidence lower limit is 2.2 for correlation and 3.0 for "pIDC" corresponding to a reduction in threshold of 0.8 units. For the 90% confidence lower limit it is 2.5 for correlation and 3.5 for the "pIDC" corresponding to a 1.0 unit reduction in threshold.

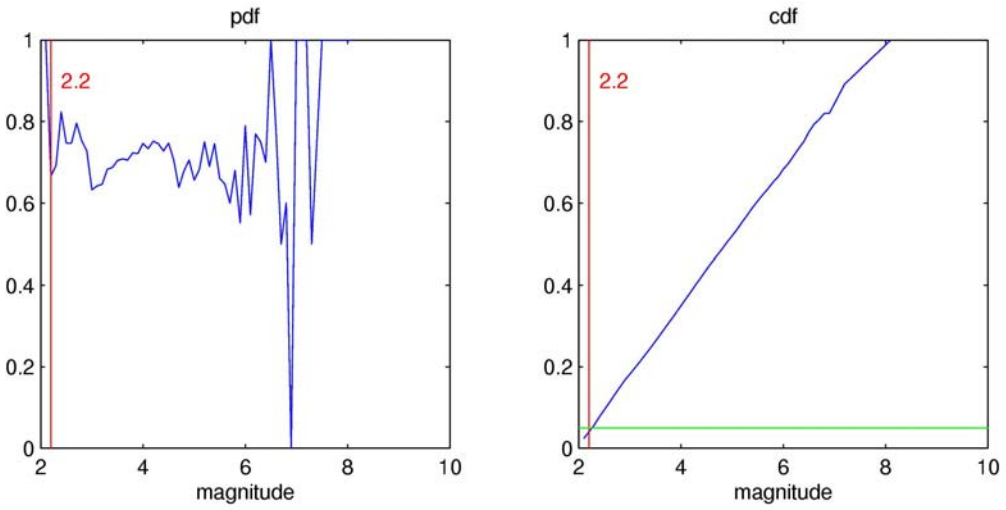

Figure 44. (left) Normalized PDF for the correlation detector for China from Figure 43. (right) Normalized CDF computed from PDF as the cumulative sum normalized to one. Magnitude 2.2 corresponds to the 95% confidence lower limit for the CDF.

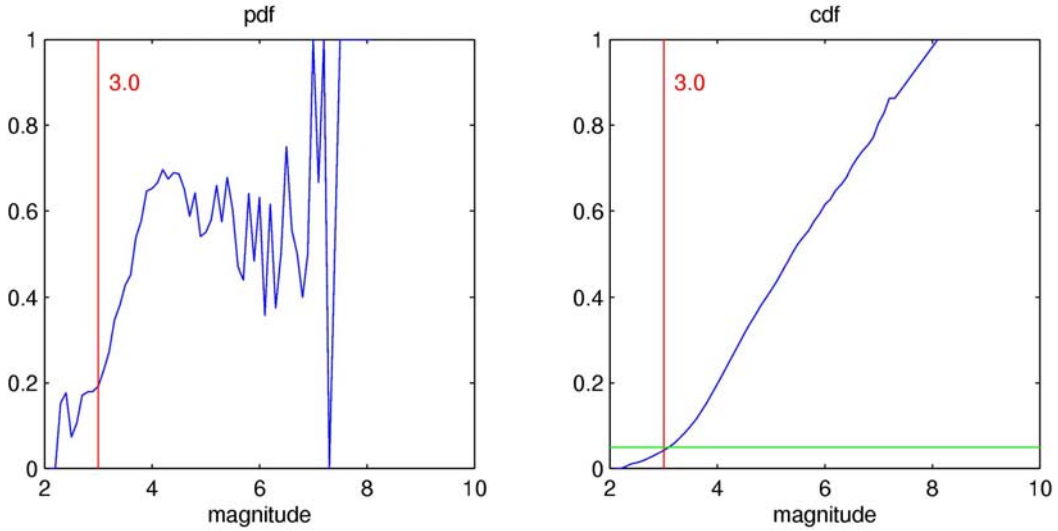

Figure 45. (left) Normalized PDF for the "pIDC" detector for China from Figure 43. (right) Normalized CDF computed from PDF as the cumulative sum normalized to one. Magnitude 3.0 corresponds to the 95% confidence lower limit for the CDF.

Table 3 summarizes and compares the results for reduction in detection thresholds using three different measures for large-scale application to China and Parkfield, California. The normalized PDF and CDF measures give more representative and intuitive results are consistent with the findings of an order of magnitude improvement from the semi-empirical synthetic runs and the smaller case study in Xiuyan, China.

Table 3. Detection threshold reduction

China	Magnitude reduction
90% confidence limit	0.3
Normalized PDF	1.7
Normalized CDF	0.9
Parkfield, California	
90% confidence limit	0.6
Normalized PDF	1.3
Normalized CDF	0.9

For the case of Parkfield the normalized measures help to reduce the impact of false alarms. For China the presence of false alarms is not so much an issue because the false alarm rate is low compared to the total number of detected events and relative to the background catalog. But for China the completeness of the catalog is an issue for lower magnitudes and the normalized measures of detection threshold appear less sensitive to that as well.

5. CONCLUSIONS

The main finding of this research is that a correlation detector can lower thresholds of detection by an order of a magnitude over standard detectors for similar events. With three component data, averaging of the scaled CC traces performs even better with a 1.3 magnitude unit reduction. Importantly, this capability is achieved with acceptably low false alarm rates as demonstrated by semi-empirical synthetic runs. Similar results of this order of magnitude improvement in detection thresholds have been demonstrated by case studies using real seismic data (Gibbons and Ringdal, 2006, Gibbons *et al.*, 2007; Schaff and Waldhauser, 2006). To estimate false alarm rates, however, synthetics must be used, or complete catalogs to lower magnitudes for denser networks on the case studies, or statistical assumptions employed (Wiechecki-Vergara *et al.*, 2001). For a SNR of 0.32 (one magnitude unit reduction compared to an STA/LTA detector), CC has a 0.5 probability of detection with 1.5 false alarms per day. Scaled CC has a slightly better false alarm rate of 1 per day for the same probability of detection. This is because it does not trigger so often on random signals of unknown origin. Scaled CC has two other benefits over CC in that it is much less dependent on window length and bandwidth which CC is highly sensitive to. In addition, the slight dependence that it does show is favorable for producing fewer false alarms compared to the dependence of CC on these parameters which can trigger many false alarms. Worthy of mention is that for nuclear

monitoring purposes missed detections are significantly more important than false alarms. Therefore tens of false alarms at single stations may be acceptable in this case. The task of associating detections to isolate single events is able to weed out many of these false alarms. Stacking of the CC traces is shown to constructively enhance the detection spikes in a manner similar to that predicted by beamforming. The detected signal has to match the 3-D particle motion and phase of the master template. Basically three independent tests have to agree giving a strong indication of a true detection and also help to weed out false alarms. As a result for a SNR of 0.32, summing the CC traces for three components has a 96% probability of detection with zero false alarms for the 36 days examined. Adding noise to the master trace as well (SNR = 1.012) in addition to the candidate trace (SNR = 0.32) on three components gives an 83% probability of detection with 1 false alarm per day.

It is important to bear in mind that these results are for similar events. Detection capability will decrease as the underlying waveform similarity decreases due to increasing inter-event separation distances, mechanisms differences, and source time function complexities. Further work has shown, however, that even semi-similar events with less than perfect waveform matches still provide useful detections (Schaff and Waldhauser, 2006).

Waveform cross correlation is a basic tool that has found applications in many fields and across many disciplines. But when it comes to observational seismology it has been used comparatively little for event detection (e. g. Gibbons & Ringdal, 2004; Gibbons & Ringdal, 2005; Gibbons & Ringdal, 2006; Gibbons et al., 2007). Shelly et al. (2006) introduced a novel, direct, scientific application of a correlation detector to identify low-frequency earthquakes in non-volcanic tremor. Still most work and practical application has concentrated on the STA/LTA technique. The greatest advances of the application of correlation in recent years have largely been seen in location estimation, improving arrival time measurements of seismic events starting with pairs of similar events or doublets and currently working up to relocating hundreds of thousands of events (e. g. Poupinet et al. 1984; Fréchet 1985; Ito 1985; Frémont & Malone 1987; Deichmann & Garcia-Fernandez 1992; Got et al. 1994; Dodge et al. 1995; Nadeau et al. 1995; Shearer 1997; Lees 1998; Rubin et al. 1999; Waldhauser et al. 1999; Waldhauser & Ellsworth 2000; Phillips 2000; Rowe et al. 2002; Schaff et al. 2002; Moriya et al. 2003; Waldhauser et al. 2004; Shearer et al. 2005; Hauksson & Shearer 2005, Waldhauser & Schaff 2007). As remarked earlier, correlation similarity thresholds much lower than needed for location purposes may be adequate for detection.

Another point worth mentioning in comparing correlation detectors to STA/LTA filters pertains to the task of associating events. Since an STA/LTA detection has little knowledge about the signal it could be from anywhere. That's why it's common to require at least three station detections to associate and locate an event before it is included bulletins. However, this poses a significant challenge for associating events where on any given day, 40% or more of the detections are not associated (National Academy of Sciences, 2002), due to the sparse nature of the International Monitoring System (IMS). In contrast, a correlation detector matches a signal with known origin time, location, and travel times to the recorded waveform and so an immediate association and hypocenter estimate can be made using just a single station. More stations just confirm a true detection and help to weed out false alarms.

We have noted how a correlation detector can aid in discriminating explosions from earthquakes for the 1997 event near Novaya Zemlya by looking for aftershocks. In this report we showed two examples of aftershocks buried in the coda of mainshocks that can be detected and weren't listed in the ABCE. Waveform cross correlation can also be used in discrimination when a signal correlates with a known explosion in the case of mining activity (Israelsson 1990; Harris 1991; Riviere-Barbier & Grant 1993) and nuclear explosions (Shearer & Astiz 1997; Thurber et al. 2001; Fisk 2002; Waldhauser et al. 2004). This issue of screening out earthquakes is very useful for large aftershock sequences such as those of the recent May 12, 2008, earthquake in China.

The case study adds to the growing body of evidence that correlation has practical application as a detector in observational seismology. It is seen that averaging of the correlation traces for three component data enhances detection spikes similar to that seen for stacking across arrays (Gibbons & Ringdal 2006). Unrelated random signals, while they may have correlation maximums that exceed the threshold on individual components, do not constructively interfere when stacked and average out to values less than the thresholds used. This is the best indication that we have made a true detection--three independent tests at a single station aligning to the nearest sample. Multiple phases (Pn, Pg, Lg) and stations (five in this case) further confirm event detections.

To extend the usefulness of a correlation detector it is desirable to see what the effect of less than perfect waveform matches is. Nearly identical repeating events represent only a fraction of the seismicity, about 10% for China (Schaff & Richards 2004), 12% for California (Waldhauser & Schaff 2007). We find that semi-similar events due to slight location, magnitude, and mechanism differences and complex source-time function histories still produce significant detection spikes for the time windows and frequency bands used. Since we will often want to detect smaller events with a larger master waveform template it is helpful to know what magnitude differences can be used before source finiteness degrades waveform similarity too much. Two examples were shown one with a 2.3 magnitude unit difference and one with a 3.3 magnitude unit difference that correlate well enough for detection (correlation maximums are significantly above background levels).

For the case study the correlation detector performs extremely well finding 90 out of 90 or 100% of the events whereas an STA/LTA detector like the pIDC employs finds 10 out of 90 or 11%. This represents a 1.3 magnitude unit reduction in detection threshold for these events. Note these results are for a sequence of similar events. Correlation detectors have remarkable sensitivity with few false alarms but require similar waveform templates. For monitoring purposes it is extremely important not to miss any detections that could be nuclear explosions even at the expense of increased false alarms. In aseismic areas there are few master events to compare to and so a clandestine nuclear test could go unnoticed if only correlation were relied upon. In seismic areas a nuclear event may also be missed if there are no previous explosions in the area to be used as master templates. Therefore it is proposed that the method be complementary and independent to the standard STA/LTA processing. Still the results are encouraging. Seismic events tend to occur in previous areas of existing seismicity whether on active faults or nuclear test sites. It is becoming more apparent that a high percentage of these events are similar enough for location and detection purposes. An analysis of 225,000 events in California

found 95% of the events correlated with at least one other event with a coefficient of 0.7 or greater at four or more stations (Schaff & Waldhauser 2005) and in China final results indicate two thirds of the 19,000 events in the ABCE correlate well enough for a detection on a much sparser network. When correlation detectors and STA/LTA filters are applied to the same network, correlation finds the significant majority of detections that standard processing detects for China (85%), for Parkfield (86%). For cross correlation to work long-term waveform archives of lots of events must be maintained. As time progresses and more events are added to the archive, the applicability of waveform cross correlation for detection, association, location, and discrimination will only improve.

Detection magnitude threshold reduction of about 1 unit holds for large scale application to 19,000 events in China and 5,000 events in Parkfield with false alarm rates of a few percent. For Parkfield the increase in number of events detected is approximately a factor of 10 like Gutenberg-Richter predicts for a magnitude unit reduction. Normalized PDF and CDF are more robust measures of detection limits, less sensitive to false alarms, and more representative of magnitude frequency distribution than unnormalized confidence limits. High resolution Parkfield locations show that the majority of the detections are for events with 1 km or less inter-event separation distances, but this is magnitude and window length dependent. Lg-waves give most detections. For some stations this may be due to larger amplitudes and durations of energy. For the two closest stations the S- and Lg- windows start nearly in the same spot so the parameters used for Lg may be the reason for the better detections -- longer windows and wider frequency bands (time-bandwidth product).

To achieve one complete magnitude unit reduction in detection threshold over a standard detector is not an easy task. It surprisingly corresponds to a signal buried in noise (invisible to the eye) at one third the noise level. Full waveform matching techniques such as those used in a correlation detector currently seem to be the only methods capable of accomplishing such a goal.

REFERENCES

Deichmann, N., and M. Garcia-Fernandez (1992). Rupture geometry from high-precision relative hypocentre locations of microearthquake ruptures, *Geophys. J. Int.* **110,** 501-5517.

Dodge, D. A., G. C. Beroza, and W. L. Ellsworth (1995). Foreshock sequence of the 1992 Landers, California earthquake and its implications for earthquake nucleation, *J. Geophys. Res.* **100**, 9,865-9,880.

Fisk, M.D., 2002. Accurate locations of nuclear explosions at the Lop Nor test site using alignment of seismograms and IKONOS satellite imagery, *Bull. Seism. Soc. Am.*, 92, 2911-2925.

Fréchet, J. (1985). Sismogenè et doublets sismiques, Thèse d'Etat, Université Scientifique et Medicale de Grenoble, 206 pp.

Freiberger, W.F., 1963. An approximation method in signal detection, *Quart. J. App. Math.*, **20**, 373-378.

Frémont, M.-J., and S. D. Malone (1987). High precision relative locations of earthquakes at Mount St. Helens, Washington, *J. Geophys. Res*. **92,** 10,233-10,236.

Gibbons, S. J., and F. Ringdal (2004). A waveform correlation procedure for detecting decoupled chemical explosions, *NORSAR Scientific Report: Semiannual Technical Summary No. 2 – 2004*, NORSAR, Kjeller, Norway, 41-50.

Gibbons, S. J., and F. Ringdal (2005). The detection of rockbursts at the Barentsburg coal mine, Spitsbergen, using waveform correlation on SPITS array data, *NORSAR Scientific Report: Semiannual Technical Summary No. 1 – 2005*, NORSAR, Kjeller, Norway, 35-48.

Gibbons, S. J., and F. Ringdal (2006). The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.* **165,** 149-166.

Gibbons, S. J., M. B. Sørensen, D. B. Harris, and F. Ringdal (2007). The detection and location of low magnitude earthquakes in northern Norway using multi-channel waveform correlation at regional distances, *Phys. Earth Planet. Inter.*, **160,** 285-309.

Got, J. -L., J. Frechet, and F. W. Klein (1994). Deep fault plane geometry inferred from multiplet relative relocation beneath the south flank of Kilauea, *J. Geophys. Res.* **99,** 15,375-15,386.

Hauksson, E., and P. Shearer (2005). Southern California hypocenter relocation with waveform cross-correlation: Part 1. results using the double-difference method, *Bull. Seism. Soc. Am.* **95,** 896-903.

Harris, D. B. (1991). A waveform correlation method for identifying quarry explosions, *Bull. Seism. Soc. Am.* **81,** 2395-2418.

Harris, D., 2006. Subspace Detectors: Theory, Lawrence Livermore National Laboratory internal report UCRL-TR-222758, 46 pp. (http://www.llnl.gov/tid/lof/documents/pdf/335299.pdf)

Harris, D. & Paik, T., 2006. Subspace detectors: efficient implementation, Lawrence Livermore National Laboratory internal report UCRL-TR-223177, 36 pp. (http://www.llnl.gov/tid/lof/ documents/pdf/336400.pdf)

Harris, D. B. (1997). Waveform correlation methods for identifying populations of calibration events, in *Proceedings of the 19th Annual Seismic Research Symposium on Monitoring a Comprehensive Test Ban Treaty September 23-25, 1997,* 604-614.

Israelsson, H., 1990. Correlation of waveforms from closely spaced regional events, *Bull. Seism. soc. Am.*, **80**(6), 2177-2193.

Ito, A. (1985). High resolution relative hypocenters of similar earthquakes by cross spectral analysis method, *J. Phys. Earth* **33,** 279-294.

Lees, J. M. (1998). Multiplet analysis at Coso geothermal, *Bull. Seis. Soc. Am.* **88**, 1127-1143.

Moriya, H., H. Niitsuma, and R. Baria (2003). Multiplet-clustering analysis reveals structural details within the seismic cloud at the Soultz Geothermal Field, France *Bull. Seis. Soc. Am.* **93**, 1606-1620.

Nadeau, R. M., W. Foxall, and T. V. McEvilly (1995). Clustering and periodic recurrence of microseismicities on the San Andreas fault at Parkfield, California, *Science* **267,** 503-507.

National Academy of Sciences, *Technical Issues Related to the Comprehensive Nuclear Test Ban Treaty*, National Academy Press, Washington, D.C., 2002.

Phillips, W. S. (2000). Precise microearthquake locations and fluid flow in the geothermal reservoir at Soultz-sous-Forêts, France, *Bull. Seism. Soc. Am.* **90,** 212-228.

Poupinet, G., W. L. Ellsworth, and J. Frechet (1984). Monitoring velocity variations in the crust using earthquake doublets: an application to the Calaveras Fault, California, *J. Geophys. Res.* **89,** 5719-5731.

Richards, P.G. & Kim, W.-Y., 1997. Testing the nuclear test-ban treaty, *Nature*, **389**, 781-782.

Riviere-Barbier F. & Grant, L., 1993. Identification and location of closely spaced mining events, *Bull. Seism. Soc. Am*, **83**, 1527-1546.

Rowe, C. A., R. C. Aster, B. Borchers, and C. J. Young (2002). An automatic, adaptive algorithm for refining phase picks in large seismic data sets, *Bull. Seism. Soc. Am.* **92,** 1660-1674.

Rubin, A. M., D. Gillard, and J.-L. Got (1999). Streaks of microearthquakes along creeping faults, *Nature*, **400,** 635-641.

Schaff, D. P., G. H. R. Bokelmann, G. C. Beroza, F. Waldhauser, and W. L. Ellsworth (2002). High resolution image of Calaveras Fault seismicity, *J. Geophys. Res.* **107,** 2186, doi:10.1029/ 2001JB000633.

Schaff, D.P., and P.G. Richards (2004a). *Lg*-wave cross correlation and double difference location: application to the 1999 Xiuyan, China, sequence, *Bull. Seismol. Soc. Am.*, **94**, 867-879.

Schaff, D. P., and P. G. Richards (2004b). Repeating seismic events in China, *Science* **303,** 1176-1178.

Schaff, D.P., Bokelmann, G.H.R., Ellsworth, W.L., Zanzerkia, E., Waldhauser, F. & Beroza, G.C., 2004. Optimizing correlation techniques for improved earthquake location, *Bull. Seism. Soc. Am.*, **94**, 705-721.

Schaff, D. P., and F. Waldhauser (2006). Improving magnitude detection thresholds using multi-event, multi-station, and multi-phase methods, *SSA Annual Meeting, April 18-22*.

Schaff, D.P., 2007. Correlation detectors applied to China seismicity, *SSA Annual Meeting*, April 11-13, Hawaii.

Schaff, D.P., 2008. Semi-empirical statistics of correlation detector performance, *Bull. Seism. Soc. Am*., *98*, 1495-1507, doi: 10.1785/0120060263.

Shearer, P. M. (1997). Improving local earthquake locations using the L1 norm and waveform cross correlation: application to the Whittier Narrows, California, aftershock sequence, *J. Geophys. Res*. **102,** 8,269-8,283.

Shearer, P. M., E. Hauksson, and G. Lin (2005). Southern California hypocenter relocation with waveform cross-correlation: Part 2. results using source-specific station terms and cluster analysis, *Bull. Seism. Soc. Am.* **95,** 904-915.

Shelly, D.R., Beroza, G.C., Ide, S. & Nakamula, S., 2006. Low-frequency earthquakes in Shikoku, Japan, and their relationship to episodic tremor and slip, *Nature*, **442**, 188-191.

Thurber, C.H., Trabant, C., Haslinger, F. & Hartog, R., 2001. Nuclear explosion locations at the Balapan, Kazakhstan, nuclear test site: the effects of high-precision arrival times and three-dimensional structure, *Phys. Earth Planet. Inter*., **123**, 283-301.

Van Trees, H.L., 1968, *Detection, Estimation and Modulation Theory*, John Wiley and Sons, New York.

Waldhauser, F., W. L. Ellsworth, and A. Cole (1999). Slip-parallel seismic lineartions along the northern Hayward fault, California, *Geophys. Res. Lett.* **26,** 3,525-3,528.

Waldhauser, F., and W.L. Ellsworth (2000). A double-difference earthquake location algorithm: method and application to the northern Hayward Fault, California, *Bull. Seism. Soc. Am.*, 90, 1,353-1,368.

Waldhauser, F., W. L. Ellsworth, D. P. Schaff, and A. Cole (2004). Streaks, multiplets, and holes: High-resolution spatio-temporal behavior of Parkfield seismicity, *Geophys. Res. Lett.* 31 L18608, doi:10.1029/2004GL020649.

Waldhauser, F., Schaff,D., Richards, P.G. &Kim,W.-Y., 2004. Lop nor revisited: underground nuclear explosion locations, 1976-1996, from double difference analysis of regional and teleseismic data, *Bull. seism. Soc. Am.*, **94**(5), 1879-1889.

Wiechecki-Vergara, S., H. L. Gray, and W. A. Woodward (2001). Statistical development in support of CTBT monitoring, *Defense Threat Reduction Agency Technical Report* **DTRA-TR-00-22,** 13-39.

List of Symbols, Abbreviations, and Acronyms

AFRL	Air Force Research Laboratory
CC	Cross correlation coefficient
SCC	Scaled cross correlation coefficient
SNR	Signal-to-noise ratio
STA/LTA	Short term average – long term average ratio